Maximum information gain in weak or continuous measurements of qdits: complementarity is not enough

Joshua Combes[1,2] and Howard M. Wiseman[1]

[1]*ARC Centre for Quantum Computation and Communication Technology, and Centre for Quantum Dynamics, Griffith University, Brisbane, 4111, Australia*

[2]*Center for Quantum Information and Control, University of New Mexico, Albuquerque, NM 87131-0001, USA*

(Dated: January 20, 2013)

To maximize average information gain for a classical measurement, all outcomes of an observation must be equally likely. The condition of equally likely outcomes may be enforced in quantum theory by ensuring that one's state ρ is maximally different, or complementary, to the measured observable. This requires the ability to perform unitary operations on the state, conditioned on the results of prior measurements. We consider the case of measurement of a component of angular momentum for a qudit (a D-dimensional system, with $D = 2J + 1$ where J is the total angular momentum). For weak or continuous-in-time (i.e. repeated weak) measurements, we show that the complementarity condition ensures an average improvement, in the rate of purification, of only 2. However, we show that by choosing the locally optimal control protocol of this type, one can attain the best possible scaling, $O(D^2)$, for the average improvement. For this protocol the acquisition of information is nearly deterministic. Finally we contrast these results with those for complementarity-based protocols in a register of qbits.

I. INTRODUCTION

In the classical world predictability is associated with intimate knowledge of a system while unpredictability implies surprise. In information theory, the surprisal of an outcome k is defined as $I(k) \equiv -\log(P_k)$, were P_k is the probability of that outcome [1]. The suprisal is an important quantity as it quantifies the amount of information one learns from an outcome of a measurement. When the average surprisal of an experiment is maximized the observer maximizes the extraction of information from the experiment.

In quantum theory the relationship between predictability and knowledge is not clear cut. Even when an observer has maximal knowledge about a quantum system (that is, the observer's state of knowledge is pure), measurement outcomes may be unpredictable. This is the case if measurements on the system are performed in a basis which is not the eigenbasis of the state. (The eigenbasis means the basis in which the operator under consideration, e.g. ρ, only has entries on the diagonal.) Nevertheless pure quantum states are the most predictable states available. It is for this reason, and others eloquently explained in Ref. [2], that we take the impurity to be a measure of information. These considerations have lead to information [2] and control [3] theoretic formulations of information–disturbance relations.

Inspired by this work, Jacobs asked the question: how quickly on average can information be extracted from a quantum system using continuous quantum measurements in a complementary basis [4]. (Other ways of quantifying the rate of information gain can also be considered [5, 6].) He found, for a widely applicable continuous measurement model, that measuring in an comple-

mentary basis led to a speed-up in the purification by a factor of two for a two level system [4]. (Practically, such a rapid purification protocol can be a primitive for a rapid state preparation protocol [6, 7].) Later this was generalized to complementary bases for D-dimensional systems. In Ref. [8] the lower bound on the speed-up for D-dimensional systems was found to be $S_{\mathrm{LB}} = (2/3)(D+1)$ when the monitoring was performed in an complementary basis. Recently in Ref. [9] the upper bound S_{UB} was proven to be bounded above by $D^2/2$ for measurement in any complementary basis.

In this paper we give an explicit construction of an complementary measurement protocol for a qudit (a D-dimensional system, with $D = 2J + 1$ where J is the total angular momentum) based on the quantum Fourier transform (QFT) which achieves a speed-up of $0.2D^2$. Further we show that transforming the state's eigenbasis so that it is complementary with respect to the measurement eigenbasis is *not* sufficient to achieve this speed-up; the speed-up can be as low as 2. In addition to this we show, numerically, that unbiased bases can achive the upper bound, saturating the inequality $S \leq D^2/2$, for D even. To define these speed-ups rigorously, it is necessary to derive rigorous bounds on information extraction rate for commuting measurements without control, so we also do that in this paper. (For the rate of commuting measurements with control see Refs. [5, 7]). Finally we consider measurements in commuting and complementary bases for a register of qbits. The complementary measurement scheme based on the QFT appears only to give a speed-up of 2 in the case of a register.

II. FURTHER BACKGROUND AND STRUCTURE OF THIS PAPER

In this section we explain the structure of our paper by reviewing the relevant prior work on the subject and relating it to the work presented in this article.

In the quantum information context, information–disturbance relations were inspired by the *epistemic* interpretation of quantum states [10] as applied to quantum cryptography [2]. The idea was that the information gathering and disturbance should be grounded with respect to an observer's state of knowledge, not some pre-existing property of the system. To formalize this intuition Fuchs and Jacobs in Ref. [2] considered something like the following game. Two observer's Alice and Bob initially agree that the state of some quantum system is ρ_0, an impure state. At some later time Charlie will perform a measurement along a randomly chosen axis and ask Alice and Bob to predict the outcome of the measurement. After repeating this many times Charlie rewards the individual who makes the most correct guesses. Alice, being clever and resourceful, has the ability to perform one unitary operation followed by a two-outcome measurement. Her aim is to refine her state of knowledge about the quantum system, thus increasing its predictability and accordingly the likelihood of receiving the money. However, Alice's morals prevent her from intentionally sabotaging Bob for the pecuniary reward. Thus Alice would like to increase the accuracy of her predictions without affecting those of Bob.

This formulation suggests (intuitively) that weak measurements are the right way to probe the system. From the point of view of the weak measurement formalism, a projective measurement corresponds to a measurement of infinite strength [2]. For finite strength measurements on a single qubit it was found that there is a non-trivial tradeoff between Alice's information gathering actions (the refinement of her predictability) and the disturbance to Bob's predictability (his state of knowledge).

For a fixed measurement rate the interval of the non-trivial tradeoff may be characterized by the angle between the Bloch vector and the measurement axis. There are two extremes of this tradeoff. Minimal disturbance and minimal *average* information gain occurs when Alice's axis is aligned with the Bloch vector, that is the measurement eigenbasis and the state's eigenbasis commute. Maximal disturbance and maximal average information gain, is attained when Alice's axis is orthogonal to the Bloch vector. This strategy corresponds to making the outcomes equi–likely, that is maximizing the *suprisal* [1] of the measurement result. In this situation the bases are said to be complementary, maximally non-commuting, or unbiased. Thus we conclude that quantum theory does not allow Alice to succeed. The ability to measure in different bases is an essential feature of quantum mechanics which provides the richness to the problem considered here.

It is easy to understand how non-commuting measurements provide maximal information gain and maximal disturbance through the following picture. Consider a mixed qbit state with its Bloch vector aligned along the $+|z\rangle$ direction. By measuring in the $|x\rangle$ basis the Bloch vector will elongate regardless of the measurement outcome. This point is illustrated in Fig. 1 (a). Furthermore both outcomes are equally likely, which means the *average* information extraction is maximal. Unfortunately the post-measurement state is not pointing in the same direction as the pre-measurement state, so it will not commute with the original state. Conversely if the measurement basis and the state are along the $|z\rangle$ direction then, due to the stochastic nature of the measurement outcomes, the average elongation of the Bloch vector will be minimal [6, 11]. In any particular measurement the impurity may increase or decrease. But the disturbance to the state will be minimal [2]. In Fig. 1 (b) we illustrate that one of the possible measurement outcomes causes the Bloch vector to shrink whist the other elongates it.
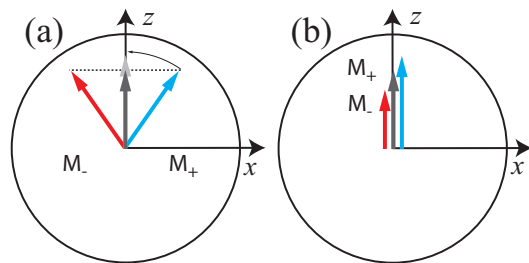

FIG. 1: A Bloch sphere depiction of the extremes of information–disturbance tradeoffs. In both figures the dark arrow represents the Bloch vector of the pre-measurement state. The arrows labeled with $M_\pm$ correspond to the post measurement states conditioned on the random outcome of a weak measurement. Figure (a) is measurement of a state in a complementary or unbiased basis, in this case the measurement axis is along the x direction. When this occurs the Bloch vector always elongates, as indicated by the light grey arrow (the plus outcome rotated for easy visual comparison). Figure (b) depicts the commuting measurement. In this case if the measurement result M_- is obtained the Bloch vector decreases in length.

While information–disturbance relations are interesting for single measurements the subject has been thoroughly explored. Less exploration has been undertaken in continuous measurement setting due to the increased difficultly of the analysis. In order to make some progress the focus on disturbance has mostly been ignored (for good reasons as we shall see). That is, the focus has been on information gathering. This restriction has led to an interesting development in the theory of continuous quantum measurements [12–14] .

To demonstrate an improvement in the extraction of information the extraction rate of a standard continuous measurement must be defined. We follow Jacobs and define the standard measurement as a measurement in a commuting basis [4]. This is sensible for a number of

reasons. Firstly measurement will rapidly destroy the coherences of any state that does not commute with the measured observable [15]. This can be understood as the measurement projecting the state to the eigenbasis of the observable. To counter this effect an adaptive measurement or closed loop feedback control is required. The second reason is a continuous measurement in a commuting eigenbasis is essentially a classical continuous measurement.

The analysis of disturbance and information gain is significantly harder for continuous measurements of D-dimensional systems [8, 16]. Minimal disturbance still occurs when measuring in the same eigenbasis as the state (a commuting basis). A full analysis of this situation would require characterization of the distribution of information extraction rates, allowing calculation of the average information extraction, median, worst case and best case. In Sec. III we rigorously derive the average purification rate for a commuting measurement of a D-dimensional system and provide an asymptotic expression. Additionally we give exact results for bounds on the spread of the distribution of purities, in Sec. III D, which were heuristically derived in previous work [16].

Now that we have defined what a standard measurement is we can explore what it means to increase the rate of extraction of information. In the context of control theory one might be interested in revealing or acquiring the value of some parameter; that is the domain of adaptive measurements [14, 17]. But more typical control objectives are to drive the evolution of the state or to stabilize the state against noise [14, 16, 18–20]. Thus disturbing the prior state is not a problem. This why disturbance is not usually considered in feedback control. What becomes important is the noise introduced by the measurement to the system. Doherty, Jacobs and, Jugnman in Ref. [3] showed, for a continuously monitored qbit, that the measurement noise is maximal when the eigenbases of the state and measurement axis are unbiased.

The study of Wiseman, Mancini and Wang in Ref. [20], explored state stabilisation in a qbit. The relevant result from Ref. [20] to the present discussion relates to the difficulty of stabilising a state which is complementary to the measurement basis. They showed that this can be achieved in the limit of infinite strength "Bayesian" feedback, but not in simpler "Markovian" feedback procedures. This is a consequence of the increased measurement noise as predicted by Doherty *et al.* [3]. Later Jacobs showed how continuously monitoring the state of a qbit in a complementary basis (assuming infinite strength feedback) leads to a doubling rate of information extraction over monitoring in a commuting basis asymptotically [4].

Recently a number of studies have examined the use of weak (and continuous) measurement in a complementarity basis in systems of dimension D greater than two. The first study showed that the lower bound on the improvement in the rate of information extraction when measuring in any complementary is $(2/3)(D+1)$ [8]. The second study showed this *speed-up* is upper bounded by a factor $D^2/2$ [9]. We summarise these results in Sec. IV A. In Secs. IV B to IV D of this paper we give an explicit construction of an unbiased basis protocol for an arbitrary finite dimensional system based on the quantum Fourier transform. We derive upper and lower bounds on the speed-up of the QFT protocol. Surprisingly the numerics show that the protocol saturates the upper bound – for even D. For completeness we mention the recent work of Ruskov et al. [21], where it was shown that monitoring a qbit in three complementary bases where they claimed a speed-up $S \sim 3$. We are currently exploring this relationship so we will not discuss it further in this article.

Motivated by some of the above investigations Jacobs and Lund argued that the task of state stabilisation is best accomplished by measurement in an complementary basis; in the regime of strong feedback [19]. This is one reason to continue investigations into maximal information extraction, foundational issues aside. In Ref. [19] the authors raised two interesting questions: first, are all complementary bases are equally good for information extraction when $D \geq 4$? Second, is measurement in a complementary basis is best for information extraction?

We reexamine these questions in Sec. V and Sec. VI. We show that complementary, i.e. equi-likelihood condition, is not sufficient to guarantee maximum information extraction. Our analysis provides an insight into the mechanism for information extraction, which is quantum mechanical backaction. From this we explain how maximum information extraction is achieved by enforcing the complementary condition and arranging the eigenvalues in phase space so measurement backaction can have maximum effect. We provide further evidence for the claim that not all bases are equally good. Because Jacobs and Lund did not examine permutations of the unbiased basis it was still unclear from their analysis whether all unbiased bases are equally powerful for information extraction. So we also consider the role of permutations in Sec. VI. Our results show that even after maximizing information extraction over permutations of the unbiased basis, not all unbiased bases are equally good at information extraction.

In Sec. VII we examine complementary measurements in a register of qbits using the QFT. We start our analysis, in Sec. VII A, with the case of commuting measurements [9]. Then, in Sec. VII B, we review the results of Ref. [9] where it was shown that the speed-up bounded by $2n(2^n-1)^{-1} \leq S \leq 2n$. In Sec. VII C we construct a protocol based on the QFT and present results which suggest the speed-up is at most 2. We conclude in Sec. VIII.

III. COMMUTING MEASUREMENTS

Later in this article we will investigate strategies that increase the rate at which information can be extracted

from quantum systems. Before these strategies are examined, we must fully characterize this rate in the absence of feedback. This will provide a benchmark to assess the performance of the strategies that aim to increase the rate of information extraction.

In this section we restrict our attention to continuous measurement (CM) of a single D-dimensional quantum system, also called a qdit, monitored by a single output channel. This output channel is monitored in such a way that it results in a diffusive quantum trajectory [12, 14]. This can be thought of as a sequence of short-duration, weak measurements. Following Fuchs and Jacobs [2] our notion of information is characterised by the impurity: $L[\rho(t)] \equiv 1 - \text{Tr}[\rho(t)]$. Accordingly we derive an exact expression for the average impurity $\langle L(t) \rangle$ and a simpler approximate expression in Sec. III A. In Sec. III B and Sec. III C we examine the validity of our approximate expression. To completely characterise the rate of information extraction from commuting measurements we look at bounding the distribution of purities in Sec. III D. Then finally we summarise the key messages in Sec. III E.

A.   Qdit - impurity analysis

The starting point for the current analysis is similar to that presented in Refs. [8, 9, 22]. We use a widely applicable model that describes the evolution of the state ρ of a quantum system due to repeated weak measurements of a dimesionless observable X. This model is described by the stochastic master equation (SME) [14, 23–25]

$$d\rho = dt\mathcal{L}_c[X]\rho \equiv 2\gamma\, dt\, \mathcal{D}[X]\rho + \sqrt{2\gamma}\, dW\, \mathcal{H}[X]\rho, \quad (1)$$

where $\mathcal{D}[A]\rho \equiv A\rho A^\dagger - \frac{1}{2}(A^\dagger A\rho + \rho A^\dagger A)$, $\mathcal{H}[A]\rho \equiv A\rho + \rho A^\dagger - \text{Tr}[(A^\dagger + A)\rho]\rho$ [14], and dW is the increment of a Wiener noise process [26]. It should be noted that we have moved to a frame that has enabled us to factor out the Hamiltonian evolution. The *measurement rate*, γ, determines the rate at which information is extracted, and thus the rate at which the system is projected onto a single eigenstate of X [16, 27–29]. (This means for times $\tau \gg \gamma^{-1}$ we may say that we have performed a projective measurement of the observable X.) The measurement result in a small time interval $[t, t+dt)$ is denoted

$$dR(t) = \sqrt{4\gamma}\langle X(t)\rangle dt + dW(t), \quad (2)$$

where dW is the same Wiener noise process that appears in Eq.(1), and $\langle X(t)\rangle = \text{Tr}[X\rho(t)]$. A quantum trajectory is a continuous–in–time description of the state conditioned on the measurement result. Thus, to specify a trajectory over the interval $[0, T]$ we must be given the measurement results over that interval. Then the Itō definition $\rho(t+dt) = \rho(t) + d\rho(t)$ may be used recursively to generate a trajectory. Clearly a single weak measurement is described by one application of this relation.

We will denote the continuous measurement record obtained by the observer integrated up until time t as

$$R(t) = \int_0^t dR(t'). \quad (3)$$

The unnormalised version of Eq. (1) is know as the *linear SME*. We denote this linear SME as [14, 30–32]

$$d\bar\rho = dt\bar{\mathcal{L}}_c[X]\bar\rho \equiv 2\gamma\, dt\, \mathcal{D}[X]\bar\rho + \sqrt{2\gamma}\, dR\, \bar{\mathcal{H}}[X]\bar\rho. \quad (4)$$

Here $\bar{\mathcal{H}}[A]\rho \equiv A\rho + \rho A^\dagger$, and the bar over ρ denotes the lack of normalization at all but the initial time. We take the initial state to be $\bar\rho = \rho(0) = \mathbf{I}/D$ where $\mathbf{I}$ is the $D \times D$ identity matrix. Because the initial state matrix commutes with X at all times this makes obtaining the solution to Eq. (4) simple. This solution is [24]

$$\bar\rho(R, t) = \exp(-4\gamma X^2 t)\exp(2\sqrt{2\gamma}XR(t))\mathbf{I}/D. \quad (5)$$

Using the linear trajectory (see Refs. [30–32]) solution, *i.e.* Eq. (5), we now explicitly calculate the form of the D-dimensional state matrix. We take the observable X to be the z component of angular momentum (represented by the operator J_z), thus $D = 2J + 1$. In matrix form the J_z operator is $J_z = \text{diag}(J, J-1, \ldots, -J+1, -J)$. Using this representation of J_z the unnormalized solution of the state evolution can be written as matrix elements exponentiated:

$$\bar\rho(R, t) = \frac{1}{D}\text{diag}\Big(e^{-4\gamma J^2 t}e^{2\sqrt{2\gamma}JR},$$
$$e^{-4\gamma(J-1)^2 t}e^{2\sqrt{2\gamma}(J-1)R},$$
$$\cdots, e^{-4\gamma(-J)^2 t}e^{-2\sqrt{2\gamma}JR}\Big). \quad (6)$$

The normalization factor can be written with a non-symmetric sum or a symmetric sum

$$\mathcal{N} \equiv \text{Tr}[\bar\rho] = \frac{1}{D}\sum_{m=0}^{2J} e^{-4\gamma(J-m)^2 t}e^{2\sqrt{2\gamma}(J-m)R} \quad (7)$$

$$= \frac{1}{D}\sum_{s=-J}^{J} e^{-4\gamma s^2 t}e^{2\sqrt{2\gamma}sR}. \quad (8)$$

For discussions later we also calculate the probability distribution of the result R and choose the symmetric sum for the normalization

$$dR\,\mathcal{P}(R) = dR\,\text{Tr}[\bar\rho(R, t)]P(R) \quad (9)$$

$$= dR\frac{1}{D\sqrt{2\pi t}}\sum_{s=-J}^{J} e^{-(R-2\sqrt{2\gamma}st)^2/2t}. \quad (10)$$

Here $P(R)$ has been called the ostensible distribution for R [14, 31]. It is given by $P(R, t) = \exp(-R^2/2t)/\sqrt{2\pi t}$. By choosing the symmetric summation for the normalization one can see that the probability distribution for R

contains D peaks centred around the values $\propto s$, where $s \in [J, J-1, \ldots, -J+1, -J]$. For times $t \gg \gamma^{-1}$, the distribution is sharply peaked about these D values. Each of these peaks has a FWHM of $2\sqrt{2\ln 2}\sqrt{t} \approx 2.35\sqrt{t}$. This seems paradoxical; one would expect that as time increases the distribution would get narrower, reflecting the fact that the observer is more confident about which eigenstate they actually have. This paradox is resolved when one realizes that the distance between the peaks increases at a rate greater than the increase in the width of the peaks. This time-dependent scaling of the distance between the peaks can be removed by changing variables to $V \equiv R/(2\sqrt{2\gamma}t)$. Under this transformation the probability distribution is

$$dV\mathcal{P}(V) = dV \frac{1}{D}\sqrt{\frac{4\gamma t}{\pi}} \sum_{s=-J}^{J} e^{-(V-s)^2/(1/4\gamma t)}. \quad (11)$$

Now the FWHM of each peak scales as $\sqrt{\ln 2/(\gamma t)} \approx 0.83/\sqrt{\gamma t}$, which clearly illustrates that the probability distribution becomes sharply peaked about the D values for long-times. In Fig. 2 the dashed curve is a plot of Eq. (11) for $D = 5$ and $t = 4\gamma^{-1}$. All of the peaks are clearly distinguishable with a measured FWHM ≈ 0.418. This agrees with the prediction given by $0.83/\sqrt{\gamma t}$.

Returning to the non-symmetric sums, i.e. Eq. (7), for the remainder of this calculation, the purity is

$$\mathrm{Tr}\left[\rho^2\right] = \frac{1}{D^2\mathcal{N}^2} \sum_{m=0}^{2J} e^{-8\gamma(J-m)^2 t} e^{16\gamma t(J-m)V}. \quad (12)$$

The impurity, for a particular value of V, is given by

$$\Lambda(V,t) = \left(1 - \frac{\sum_{m=0}^{2J} e^{-8\gamma(J-m)^2 t} e^{16\gamma t(J-m)V}}{D^2\mathcal{N}^2}\right) \quad (13)$$

$$= \left(\frac{D^2\mathcal{N}^2 - \sum_{m=0}^{2J} e^{-8\gamma(J-m)^2 t} e^{16\gamma t(J-m)V}}{D^2\mathcal{N}^2}\right). \quad (14)$$

The solid curve in Fig. 2 is a plot of Eq. (14) for $D = 5$ and $t = 4\gamma^{-1}$. To find the average value of impurity for a continuous measurement of J_z one must integrate this function over the scaled measurement record V, weighted by $P(V)$ [8]:

$$\langle L(t)\rangle = \int_{-\infty}^{\infty} \Lambda(t, V)\mathcal{N}(V)P(V)dV. \quad (15)$$

When using this result to calculate the measurement-induced evolution of the mean linear entropy one should use the kernel in Eq. (14) rather than the kernel in Eq. (13). This is because it will reduce the numerical error that occurs when subtracting from one a number very close to one.

To calculate this integral one must perform numerical integration. It would be useful to have a fully analytic
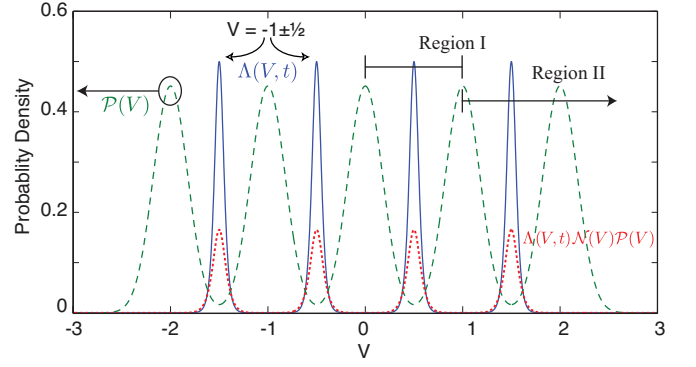

FIG. 2: The dashed curve is the probability distribution $\mathcal{P}(V)$, Eq. (11); the solid line is $\Lambda(V, t)$, Eq. (14); shown for $D = 5$ at $t = 4\gamma^{-1}$. The dotted curve is $\Lambda(V, t)\mathcal{N}(V)P(V)$ multiplied by 20 (so that it can be viewed on the same scale). Also plotted are regions for the two eigenvalue expression for the impurity integral.

expression for $\langle L(t)\rangle$ in the D-dimensional case. Now we will make some approximations to obtain a simple analytic expression for Eq. (15). We will show that the approximations allow one to place a bound on the full expression. Then we will confirm our approximations with numerical simulations in Sec. III B and Sec. III C.

The first approximation involves truncating the state matrix to two eigenvalues and then renormalizing. We call this the *two eigenvalue approximation*. The motivation for the truncation stems from the following observations. In the long-time limit ($t \gg \gamma^{-1}$) the true probability distribution, $\mathcal{P}(V) = \mathcal{N}(V)P(V)$, is sharply peaked in D places; see Fig. 2. Also in this limit one finds that $\Lambda(V, t)$, Eq. (14), is sharply peaked in $(D-1)$ places; the peaks are between peaks of $\mathcal{P}(V)$. Wiseman and Ralph have pointed out, for a qbit, that this arrangement leads to poorly purifying trajectories dominating $\langle L(t)\rangle$ [6]. In a qbit the poorly purifying trajectories are those trajectories with eigenvalues of the same magnitude ($V(t) \approx 0$) in the long-time limit. Physically this means that the measurement has not projected, or the filter has not decided, if the state is in $z = \pm\frac{1}{2}$. In qdits there are $(D-1)$ regions like this, where $V(t) \approx r$ for $r \in [J - \frac{1}{2}, J - \frac{3}{2}, \ldots, J + \frac{3}{2}, -J + \frac{1}{2}]$. In any of these regions only two eigenvalues contribute significantly to the integrand. Consequently it is reasonable to truncate the state to the two eigenvalues in a particular region and renormalize.

The effect of the two eigenvalue approximation is twofold: it increases the purity of the state, and makes it possible to derive a fully analytic expression for the impurity. The eigenvalues of truncated state matrix are $(\lambda_0, \lambda_1)/(\lambda_0 + \lambda_1)$[58], where λ_0 is the largest eigenvalue of ρ and $\lambda_a > \lambda_b$ when $a < b$. It is easy to show that ρ is majorized [33] by ρ_2 (that is, $\rho \prec \rho_2$), meaning that the original state is more mixed than our two eigenvalue approximation to it. From the fact that the purity is

Schur-convex [34] it follows that $\text{Tr}\left[\rho^2\right] \leq \text{Tr}\left[\rho_2^2\right]$. This means the impurity of the two eigenvalue approximation is a lower bound on the true impurity: $L(\rho) \geq L(\rho_2)$. In the long-time limit it is reasonable to expect that $L(\rho) \approx L(\rho_2)$ from the arguments above.

Now we calculate the impurity of the truncated state matrix $\langle L_2(t)\rangle$. We split the integral in Eq. (15) into regions such that the same two eigenvalues are the largest two eigenvalues for all values of V in each region. The integration is then performed in each of these regions and then the regions are summed. Figure (2) illustrates how the regions should be split for $D = 5$. After a coordinate transform (for example in region I it would be $V' = V - 1/2$) the integral in region one becomes

$$\mathfrak{R}_{\mathrm{I}} = \frac{e^{-\gamma t}}{D}\sqrt{\frac{4\pi t}{\pi}}\int_{-\frac{1}{2}}^{\frac{1}{2}}\frac{e^{-4\gamma t V'^2}}{\cosh\left(4\gamma t V'\right)}dV'. \quad (16)$$

After a similar transformation the integral in region II becomes

$$\mathfrak{R}_{\mathrm{II}} = \frac{e^{-\gamma t}}{D}\sqrt{\frac{4\pi t}{\pi}}\int_{-\frac{1}{2}}^{\infty}\frac{e^{-4\gamma t V'^2}}{\cosh\left(4\gamma t V'\right)}dV'. \quad (17)$$

The total integral is thus

$$\langle L_2(t)\rangle = 2\mathfrak{R}_{\mathrm{II}} + (D-3)\mathfrak{R}_{\mathrm{I}}. \quad (18)$$

For $t \gg \gamma^{-1}$ the integrands become sharply peaked about $V' = 0$ with negligible contributions to the integrals from the region outside the interval $V' \in [-\frac{1}{2}, \frac{1}{2}]$, so that $\mathfrak{R}_{\mathrm{II}} \simeq \mathfrak{R}_{\mathrm{I}}$. In this limit we have

$$\langle L_2(t)\rangle \simeq (D-1)\frac{e^{-\gamma t}}{D}\sqrt{\frac{4\pi t}{\pi}}\int_{-\frac{1}{2}}^{\frac{1}{2}}\frac{e^{-4\gamma t V'^2}}{\cosh\left(4\gamma V'\right)}dV'. \quad (19)$$

The final approximation is arrived at by noting that in the long-time limit the distribution in the numerator is broad compared to the distribution in the numerator for $t \gg \gamma^{-1}$. Thus we can make the approximation: $\int_{-\infty}^{\infty}dV e^{-4\gamma t V'^2}/\cosh(\sqrt{2\gamma}V) \approx \int_{-\infty}^{\infty}dV/\cosh(\sqrt{2\gamma}V) = \pi/\sqrt{2\gamma}$. The final expression for the impurity is thus

$$\langle L_2(t)\rangle_{\mathrm{LT}} = \frac{2(D-1)}{D}\frac{\pi e^{-\gamma t}}{\sqrt{16\gamma t\pi}}. \quad (20)$$

This is the analytic expression for $\langle L(t)\rangle$ that we set out to find, and is the key result of this section.

Because previous work has used the exact qbit results we write it out for future reference:

$$\langle L(t)\rangle_{\mathrm{qbit}} = \frac{e^{-\gamma t}}{2}\sqrt{\frac{4\pi t}{\pi}}\int_{-\infty}^{\infty}\frac{e^{-4\gamma t V'^2}}{\cosh\left(4\gamma t V'\right)}dV'. \quad (21)$$

The long time limit of this expression is

$$\langle L_2(t)\rangle_{\mathrm{qbit-LT}} = \frac{\pi e^{-\gamma t}}{\sqrt{16\gamma t\pi}}. \quad (22)$$

As could reasonably be expected for a qdit under the two-eigenvalue approximation (effectively a two-level system), the resulting average impurity is proportional to the qbit impurity with the proportionality depending on the qdit dimensionality. In Sec. III B we examine the accuracy and validity of this approximation.

B. Comparison with numerics

In Fig. 3 (a) one can see that for $t \gtrsim \gamma^{-1}$ the impurity, found by numerically evaluating Eq. (15), is better approximated by Eq. (18) than Eq. (21). Moreover, for all times the qbit impurity Eq. (21) is a lower bound on the true impurity. However, when calculating the speed-up one is interested in times where $\gamma t \gg 1$, which are plotted in Fig. 3 (b).
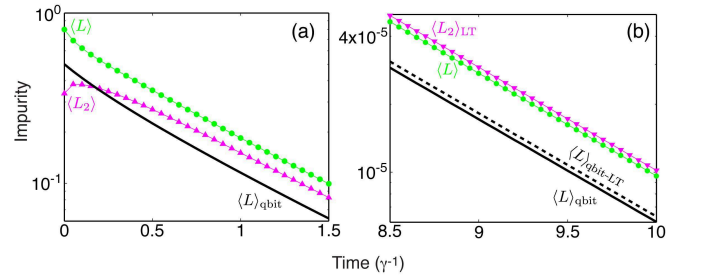

FIG. 3: A comparison of approximations to the evolution of $\langle L\rangle$, for $D = 3$. (a) (starting from the upper-most curve): $\langle L\rangle$ i.e. Eq. (15) shown by green circles; $\langle L_2\rangle$ i.e. Eq. (18) with the integrals evaluated numerically, shown by magenta upward pointing triangles; $\langle L\rangle_{\mathrm{qbit}}$ i.e. Eq. (21) shown by solid black line. (b) (starting from the upper-most curve): $\langle L_2\rangle_{\mathrm{LT}}$ i.e. Eq. (20) shown by magenta downward pointing triangles; $\langle L\rangle$ i.e. Eq. (15) shown by green circles; $\langle L\rangle_{\mathrm{qbit-LT}}$ i.e. Eq. (22) shown by dashed black line; $\langle L\rangle_{\mathrm{qbit}}$ i.e. Eq. (21) shown by solid black line. Eq. (18) is not plotted in (b) because on this scale it is indistinguishable from $\langle L\rangle$.

Eq. (18) is not plotted in Fig. 3 (b) because on this scale it is practically indistinguishable from $\langle L\rangle$ at long-times (it approaches $\langle L\rangle$ from below). But $\langle L_2(t)\rangle_{\mathrm{LT}}$, that is Eq. (20), approaches $\langle L\rangle$ from above. Thus the approximations made to obtain Eq. (20) from Eq. (18) increase the impurity. This effect is easily understood once one realizes that the approximations amount to: throwing away parts of the integral Eq. (19); and over-estimating the integral by approximating $\exp\left(-4\gamma t V'^2\right) = 1$. The next closest curve is that of the qbit long-time limit expression Eq. (22), and below that is the qbit exact result Eq. (21).

From these curves we infer the following: the long-time expression for the two eigenvalue approximation, $\langle L_2(t)\rangle_{\mathrm{LT}}$, is closer to $\langle L\rangle$ than the qbit long-time limit; $\langle L_2(t)\rangle_{\mathrm{LT}}$ is approaching $\langle L\rangle$ in the same way $\langle L(t)\rangle_{\mathrm{qbit-LT}}$ is approaching $\langle L(t)\rangle_{\mathrm{qbit}}$. This is important, as previously in the literature it has been com-

mon to use the qbit long-time limit expression as a lower bound for $\langle L \rangle$ any dimension [5, 8]. It is clear that for all times $t \gg \gamma^{-1}$, $\langle L_2(t) \rangle_{\mathrm{LT}}$ is a better approximation to $\langle L \rangle$ than the qbit expression. Of course asymptotically the expressions decay at the same rate. Therefore there is nothing in this work to suggest that the speed-up calculated in previous works are incorrect.

C. Stochastic simulations

We now compare the analytic solutions from Section III A and the exact numerics from Section III B, to stochastic non-linear trajectory simulations of Eq. (1). We used an Euler integration method with the following parameters: $\delta t = 1 \times 10^{-4} \gamma^{-1}$; $D = 5$; ensemble size $= 20$. The ensemble size is small because we will plot all of the trajectories so that we may gain some qualitative understanding from them.

Figure 4 (a) shows that the L of all of the trajectories is bounded from above by $\frac{1}{2}$ for $t \gg \gamma^{-1}$. Intuitively this effect can be understood from the two eigenvalue approximation; in the long-time limit a poorly purifying trajectory has two eigenvalues and thus these trajectories are upper bounded by $L = \frac{1}{2}$. A more rigorous explanation is given in Sec. III D. On this Figure we also plot the numerically calculated ensemble average impurity and the linear trajectories solution, Eq. (15).

There is some difficulty in obtaining convergence for $\langle L \rangle$ in the stochastic simulations, even with very large ensemble sizes. The convergence problem is due to a small number of poorly purifying trajectories first noted for a qbit in Ref. [6]. These poorly purifying trajectories are evident in Fig. 4 (b) in that many trajectories touch the dot-dashed line at $L = \frac{1}{2}$. For moderate ensemble sizes the whole region of possible impurities is filled by the ensemble even at $t = 2\gamma^{-1}$. As noted above the major contributions to $\langle L \rangle$ are from the regions where $V = r$ (see Fig. 2); that is, for results V that are as far as possible from the most likely values of V [the peaks of $\mathcal{P}(V)$]. Thankfully one can use Eq. (15) to calculate the evolution of the mean impurity. The advantages of this method are two fold: the computation time is greatly reduced, and it is exact, to numerical precision.

Figure 4 (b) shows that many trajectories also touch the dashed line corresponding to a scaling $L \sim \exp(-4\gamma t)$. The reason for this clustering, and that at $L = \frac{1}{2}$, will be explained in Sec. III D.

D. The distribution of impurities

In Ref. [16] Stockton, van Handel, and Mabuchi gave reasonable but non-rigorous arguments for a bound [their Eq. (41)] on the degree of spreading of trajectories (that is, on the width of the distribution of trajectories) for the model of Eq. (1). We now use linear trajectories to rigorously derive the bounds that they found. In all cases
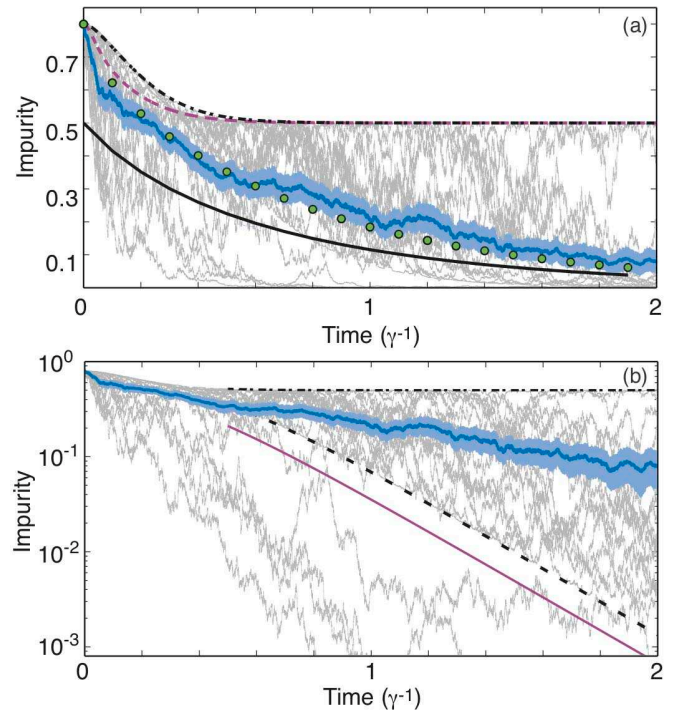

FIG. 4: Short time impurity dynamics for $D = 5$. We plot 20 impurity trajectories (grey); the ensemble average of the trajectories (blue); the error bars for the ensemble average, $\pm\sigma_{\mathrm{mean}}$, are smeared around the mean (light blue). Time is measured in units of γ^{-1}. Fig. (a) Linear abscissa. The dashed and dot-dashed lines are approximate upper bounds on the L trajectories; see Sec. III D for further explanation. The solid black line is the linear trajectories solution for the qbit, that is Eq. (21). The circles are $\langle L \rangle$ for $D = 5$, i.e. Eq. (15). Fig. (b) the impurity is plotted on a logarithmic axis; this helps to see the approximately exponential decrease in impurity predicted by Eq. (20). The approximate upper bound on the spread of L trajectories is the dot dash line. The dashed line corresponds to the most likely trajectories. Below the solid (magenta) line there is only $1/D$ of the ensemble of trajectories at long-times. These lines are explained in Sec. III D.

it is easy to obtain an analytic expression, but the exact form of the expression depends on D. Therefore we quote only the *long-time* asymptotic scaling of the bounds.

Central to our explanation of the bounds of the distribution of trajectories is the intuition gained by examining Fig. 5, where we have replotted portions of Fig. 2 on two separate axes. In Fig. 5 (a) $\mathcal{P}(V, t)$ is plotted. This probability density plot is significant in the current analysis because it shows which records, V, are likely (the peaks of $\mathcal{P}$) and which are unlikely (the troughs). In Fig. 5 (b) the kernel $\Lambda(V, t)$, that is Eq. (14), is plotted on a log scale. The key features of this plot are: 1) the peaks at $V = -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}$ which represent very impure trajectories and correspond to unlikely records, and 2) the troughs at $V = -1, 0, 1$ coincide with some of the purest trajectories and correspond to minima of $\mathcal{P}(V, t)$. We will

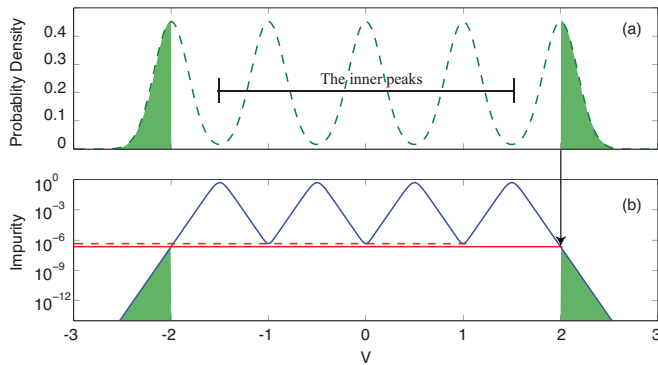elaborate on other relevant features of this graph below as each bound is explained.

FIG. 5: (a) probability distribution $\mathcal{P}(V,t)$, Eq. (11), for $D = 5$ and $t = 4\gamma^{-1}$. The shaded regions of the probability distribution contain $1/D$ of the total probability at long-times. (b) the solid blue line is $\Lambda(V,t)$, Eq. (14), plotted with logarithmic scale on the abscissa for $D = 5$ and $t = 4\gamma^{-1}$. The red dashed line is a projection onto the abscissa of the minima of $\Lambda(V,t)$ at $V = -1, 0, 1$. The red solid line is a projection onto the abscissa of the $\Lambda(V,t)$ evaluated at $V = \pm 2$; notice that this projection is purer than the dashed line. The shaded regions in this figure depict the impurity of records that have $|V| \geq J$. As these regions concur with the regions in Fig. (a) they can be interposed as saying that only a proportion $1/D$ of the trajectories will have impurities smaller than those given by the projection $\Lambda(\pm J, t)$.

1. *The upper bound:* This is obtained by substituting $V(t) = r$ (recall that $r \in [J-\frac{1}{2}, J-\frac{3}{2}, \ldots, J+\frac{3}{2}, -J+\frac{1}{2}]$) into Eq. (6), normalizing, and then calculating the impurity. This procedure gives the upper bound because these records correspond to the most impure trajectories, that is peaks in Fig. 5 (b). Because this procedure also corresponds to a a worst case scenario of the two eigenvalue approximation it can also be interpreted as the filter being unable to decided between the eigenvalues $\lambda_{r-\frac{1}{2}}$ and $\lambda_{r+\frac{1}{2}}$. For $D = 5$ the records of interest correspond to $V = -\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{3}{2}$ (see Fig. 5 and Fig. 2). In the long-time limit all solutions give

$$L_{\mathrm{LT}} = \tfrac{1}{2}, \qquad (23)$$

which is also the bound predicted by Stockton and coworkers [16]. For small times there is some variation between the bounds given by $V = r = \pm\frac{1}{2}$ and $V = r = \pm\frac{3}{2}$. This is because $\mathcal{P}(V)$ is broad at short-times so that more eigenvalues contribute to the impurity when for example $V = \frac{1}{2}$. Conversely when $V = \frac{3}{2}$ the is only one eigenvalue to the left of this record. Accordingly it is reasonable to expect that at short-times the bound given by $V = \pm\frac{1}{2}$ will produce a bound that is more impure than the bound given by $V = \pm\frac{3}{2}$. This behaviour is confirmed in Fig. 4 (a), where the dashed line is $V = \pm\frac{3}{2}$ and the dot dashed line is $V = \pm\frac{1}{2}$.

2. *A pseudo lower bound:* When $V(t) = s$ for $s \in [J-1, \ldots, -J+1]$, the record corresponds to the inner peaks of the probability distribution, as depicted in Fig. 5 (a) (the peaks at $\pm J$ are *not* included in this analysis). These peaks of $\mathcal{P}(V)$ are the most likely records and coincide with some of the purest trajectories, the minima in Fig. 5 (b). For $D = 5$ the peaks of interest are those corresponding to $V = -1, 0, 1$ in Fig. 5. By substituting $V(t) = s$ into Eq. (6), normalizing, and then calculating the impurity, one obtains the lower bound without any approximations. These solutions can be obtained for any dimension D. By making a two-eigenvalue approximation one can solve the evolution of the impurity analytically giving

$$L_{\mathrm{LT}} \sim \exp(-4\gamma t). \qquad (24)$$

Equation (24) corresponds to the lower bound quoted by Stockton and coworkers [16]. In their case, their bound appeared to be a true lower bound as their initial state was a coherent spin state (a collection of $D/2$ spin one half particles where $D = 2J + 1$ with mean spin vector, J, aligned along the x-axis in this case). This has a small population in $|\pm J\rangle$, and it is therefore unlikely in their case that $|V| > |J|$. For our analysis we see clearly that the bound is not a true lower bound (see Fig. 4 (a)), because a maximally mixed state has equal populations in all eigenstates. However, because these records $V(t) = j$ are the most likely records many of the trajectories touch or cluster around this bound. This will be further explored in point 4. below.

3. *Physically likely trajectories:* In our case there is no lower bound, although a trajectory that purifies infinitely fast is infinitely unlikely [59]. It is possible to give a natural bound on the physically likely trajectories based on the probability distribution $\mathcal{P}(V)$. In the lower bound section above we did not include the peaks of the probability distribution found at $V = \pm J$. It is clear from Fig. 5 (b) that there is no corresponding minimum of the kernel $\Lambda(V,t)$ in this region. When $\Lambda(V,t)$ is evaluated at $V(t) = \pm J$ one finds that it gives impurities smaller (see the solid red line in Fig. 5 (b)) than those which correspond to the inner peaks, [the dashed red line in Fig. 5 (b)]. From the shaded regions in Fig. 5 (a) it is apparent that only $1/D$ of the probability density is found for $|V| \geq J$. This means that, in the long-time limit, only a proportion $1/D$ of the trajectories will have purities greater than the bound we now present. Using the linear trajectory solution with $V(t) = \pm J$ and solving for the impurity one obtains the bound. By making a two eigenvalue approximation one finds that the bound scales asymptotically as

$$L_{\mathrm{LT}} \sim \exp(-4\gamma t). \qquad (25)$$

The scaling in Eq. (25) and Eq. (24) is precisely the same. The difference in the bounds is only apparent when the calculation is exact; see Fig. 4 (b). The dashed line (black) is Eq. (24), while the thin solid line (magenta) is

Eq. (25) (also see Fig. 6). Obviously for a qbit ($D = 2$) the bound given by $V = \pm\frac{1}{2}$ corresponds to the median of the distribution of impurites, denoted by $\wp(L, t)dL$, at all times, and the mode at long-times.

4. *Distribution of Impurities:* The distribution of impurites can be calculated analytically for $D = 2$ [6]. However we will change variables to the log-impurity so that we may clearly see the featured discussed above. The required change of variables from Eq. (24) in Ref. [6] is $\ell \equiv \log_{10} L = \log_{10}\left[\frac{1}{2}(1 - z^2)\right]$ so that $z = \sqrt{1 - 2 \times 10^\ell}$. The distribution of log-impurites is thus

$$d\ell\wp(\ell, t) = d\ell \frac{e^{-\gamma t}\ln(10)\cosh\left[\operatorname{arctanh}^2(\sqrt{1 - 2 \times 10^\ell})\right]}{2\sqrt{8\pi\gamma t}\sqrt{1 - 2 \times 10^\ell}}$$
$$\times \exp\left(-\operatorname{arctanh}^2(\sqrt{1 - 2 \times 10^\ell})/4\gamma t\right). (26)$$

In Fig. 6 this distribution is plotted at $t = 2\gamma^{-1}$ (the red dashed curve). The peak of this distribution at $\ell = -0.301$ becomes less prominent at long-times. This is when, as Wiseman and Ralph pointed out [6], considering the average log-impurity is a good way to find an approximation to the mean time to a fixed L.

The poorly purifying trajectories affect $\langle L \rangle$ and a large number of these trajectories cluster around the upper bound (as seen in Fig. 4). This motivates the consideration of another measure of mixedness which deemphasises the peaks of the distribution [5, 6]. However, in this article, we will remain focused on $\langle L \rangle$.
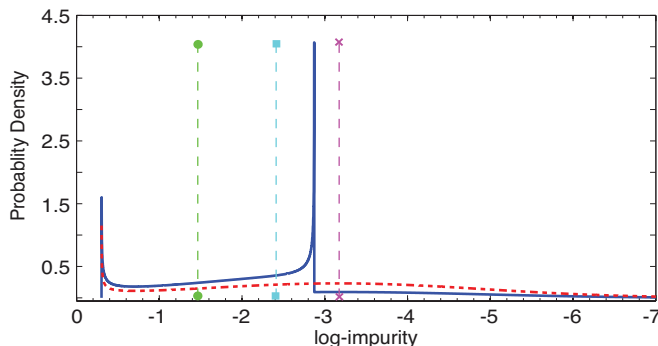

FIG. 6: The distribution of log-impurities at $t = 2\gamma^{-1}$. The blue solid curve is the distribution for $D = 5$. The red dashed curve is for $D = 2$, that is Eq. (26). The line with the circles is $\log_{10}(\langle L \rangle) = -1.46$ at $t = 2\gamma^{-1}$, according to Eq. (15); the line with the squares is $\langle \ell \rangle = -2.41$ at $t = 2\gamma^{-1}$. Only a proportion $1/D$ of the trajectories have lower purity than (to the right of) the line with the crosses (plotted at $\ell = -3.1736$) as explained in the main text. The two other features on the graph at $\ell = -0.3010$ and $\ell = -2.87$ correspond to the upper and the pseudo-lower bounds on the distribution of impurities described in the main text.

For $D > 2$ the distribution of ℓ can be numerically calculated from the probability distribution $\mathcal{P}(V)$ and $\Lambda(V, t)$ as follows. First one splits the distribution $\Lambda(V, t)$ into regions $\{(-\infty, -j), [-j, -j + \frac{1}{2}), \ldots, [j, \infty)\}$. These regions are chosen because ℓ is always monotonically increasing or decreasing in these regions; see Fig. 5 (b). This is required to perform the numerical interpolation required by the following procedure. In each region one discretizes the range of ℓ and then finds the values of V that correspond to this ℓ. The probability distribution $\mathcal{P}(V)$ is then integrated over this region to give $d\ell\wp(\ell, t)$. This is then summed over all regions.

In Fig. 6 the solid curve is the probability distribution for $D = 5$ at $t = 2\gamma^{-1}$; this distribution was obtained using the above method. The upper bound and pseudo-lower bound described in the text above are quite apparent in distribution; they correspond to the two sharp features at $\ell = -0.301$ and $\ell = -2.87$. It is also apparent that $\langle \ell \rangle$ is a more faithful central tendency measure than $\log_{10}\langle L \rangle$ as the line corresponding to $\langle \ell \rangle$ is closer to the bulk of the distribution.

E. Summary

We conclude our study of information extraction from commuting measurements by summarising the key points. It is possible to characterise the average amount of information extracted using the average impurity. We found a simple expression for the average impurity, $\langle L_2(t) \rangle_{\mathrm{LT}} \sim \exp(-\gamma t)$, which is valid when $t \gg \gamma^{-1}$. Additionally we found that in this asymptotic limit a portion $(1 - 1/D)$ of the trajectories would be bounded by $c\exp(-4\gamma t) \leq L \leq \frac{1}{2}$ for a constant c that we can determine for a particular D. It is the existence of rare but poorly purifying trajectories which explains the upper bound of $\frac{1}{2}$, and which explains the difference between the scaling of the mean and that of the lower bound, $c\exp(-4\gamma t)$. Indeed, other measures of central tendency, such as the median, or the exponential of the mean-log, exhibit the same scaling as the lower bound, $\sim \exp(-4\gamma t)$, which is also reflected in the mean time to attain a given purity [7].

Despite its limitations, we have used $\langle L \rangle$, which scales as $\exp(-\gamma t)$, as our measure of average information in the main text of the paper. Here we briefly comment on what would change if we were to have used a more faithful measure of central tendency such as the median. In Sec. IV B we found that the spread of trajectories for the QFT protocol was small for sufficiently frequent feedback. This feature is true in general for measurement in an unbiased basis [9]. The main consequence of this, for complementary continuous measurements, is the mean impurity is approximately equal to the median impurity. Thus we can directly compare the results obtained for $\langle L \rangle$ in the main part of the paper to the scaling for the median impurity, $\sim \exp(-4\gamma t)$, for commuting measurement. The result, obviously, is a diminution in the speed-up offered by the latter by a factor of four: $\frac{1}{6}(D + 1) \leq S^{\mathrm{median}} \leq \frac{1}{8}D^2$. For small systems ($D = 2$ or 3) this implies a slow-down, or at best a modest speed-up. But for large D the scaling with D implies that using

feedback to construct suitable measurements in an unbiassed basis will beat the no-feedback protocol even in terms of the median impurity or mean time.

IV. INFORMATION ACQUISITION IN DIMENSION D AND ITS RELATIONSHIP TO COMPLEMENTARITY

A. Previous results

In this subsection we briefly summarise the results of previous studies on information acquisition by measurement in a complementary basis. The original articles on the subject are found in Refs. [2, 4, 8, 9]. The basic idea is to use quantum feedback control to continuously keep the eigenbasis of the state complementary to the measurement basis. By analogy with the intuition presented in Fig. 1 it is hoped that this procedure will enhance information extraction.

It is easy to show, from Eq. (1), using Itō calculus that the information acquired [2], as characterised by the change in impurity, about the state due to measurement of infinitesimal duration is [9]

$$
\begin{aligned}
dL = {} & -8\gamma dt\{\operatorname{Tr}\left[\rho X\rho X\right] - 2\operatorname{Tr}\left[X\rho\right]\operatorname{Tr}\left[X\rho^2\right] \\
& + \operatorname{Tr}\left[\rho X\right]^2\operatorname{Tr}\left[\rho^2\right]\} \\
& -4\sqrt{2\gamma}dw\{\operatorname{Tr}\left[X\rho^2\right] - \operatorname{Tr}\left[\rho X\right]\operatorname{Tr}\left[\rho^2\right]\}.
\end{aligned}
\tag{27}
$$

We wish to study the effect of measuring in a complementary basis. In order to measure in a complementary basis through out the measurement process one must use quantum feedback control [14] to continually adjust the basis. We will not labour on those details here as they have been adequately discussed before [4, 8, 9, 14]. The required transformation of the measurement basis is $X(t) \to \check{X}(t)$ such that $|\langle\check{x}|i\rangle| = 1/\sqrt{D}$, for all $\check{x}$ and i where $|i\rangle$ is an eigenstate of ρ and $|\check{x}\rangle$ is an eigenstate of $\check{X}$. Replacing all the X's by $\check{X}$ in Eq. (27) considerably simplifies the expression to [8, 9]

$$
dL = -8\gamma\operatorname{Tr}\left[\check{X}\rho\check{X}\rho\right]dt,
\tag{28}
$$

or equivalently

$$
dL = -8\gamma dt\sum_{r,c}|\check{X}_{r,c}|^2\lambda_r\lambda_c.
\tag{29}
$$

Observe that the unbiased condition which simplified Eq. (27) to Eq. (28) leaves a permutational degree of freedom in Eqs. (28) and (29) because a permutation of the eigenstates of ρ does not affect $|\langle\check{x}|i\rangle| = 1/\sqrt{D}$. This means we should optimise over permutations in the expression $dL = -8\gamma\operatorname{Tr}\left[\check{X}P_m^\dagger\rho P_m\check{X}P_m^\dagger\rho P_m\right]dt$, where P_m is a permutation matrix, to maximize the decrease in impurity. To obtain a lower bound $|dL|$ of the optimal permutation the we average over all possible permutations: $dL = -8\gamma\sum_m\operatorname{Tr}\left[\check{X}P_m^\dagger\rho P_m\check{X}P_m^\dagger\rho P_m\right]dt$. The average

gives a lower bound because the sum of a sequence is always less than or equal to the greatest term in the sum.

It was shown in Refs. [8, 9] that a tight lower bound on Eq. (28), for the optimal permutation, is

$$
dL \leq -\frac{2}{3}\gamma t(D+1)L(t).
\tag{30}
$$

While the upper bound was recently found to be

$$
dL \geq -\frac{1}{2}\gamma tD^2 L(t),
\tag{31}
$$

for $D \gg 1$ [9]. Equations (30) and (31) hold for measurement in *any* complementary basis. Using these bounds one may compare the time it takes for a commuting measurement (*i.e.* Eq. (20)) to extract a certain amount of information (i.e. to attain $\langle L\rangle = \epsilon$) to how long a complementary measurement may take to extract that same amount of information. That is we equate Eq. (20) and the solutions of Eqs. (30) and (31), and solve for the ratio of $t_{\text{complementary}}$ and t_{commute}. We call this ratio the speed-up in information acquisition. For $t_{\text{commute}} \gg \gamma^{-1}$ the speed-up is bound by

$$
\frac{2}{3}(D+1) \leq S \leq \frac{D^2}{2},
\tag{32}
$$

for $D \gg 1$ [9].

While these results show the usefulness of feedback in rapid purification, there are some important issues that have not yet been addressed. First, does an explicit construction of such a complementary measurement strategy exist? Second, is actual performance of such a strategy $O(D)$ or $O(D^2)$? Third, what is the mechanism of the speed-up? The first and second questions are answered in Sec. IV B and Sec. IV C respectively. We defer the third question until Sec. V.

B. Complementarity via quantum Fourier transform for $D = 2, 3, 4$

In this subsection we explore implementing a continuous complementary measurement in systems of dimension three and four. There are many unbiased bases, so to make our analysis concrete we must choose one. It has been noted elsewhere [19] that for $D > 3$ not all unbiased bases are equally good at reducing the impurity. Nevertheless we choose a particular complementary, the one generated by the quantum Fourier transform (QFT) of the logical basis. The construction of the continuous complementary measurement then consists of: calculationally diagonalizing the state matrix; calculationally ordering (permuting) those eigenvalues to maximize $|\langle dL\rangle|$ after a measurement; and applying the appropriate permutation and the QFT unitary to the state. This whole procedure is an example of quantum feedback control [14]. The D-dimensional QFT can be represented by

the matrix

$$T = \frac{1}{\sqrt{D}} \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & q & q^2 & \dots & q^{1(D-1)} \\ 1 & q^2 & q^4 & \dots & q^{2(D-1)} \\ \vdots & \vdots & \vdots & \ddots & \\ 1 & q^{(D-1)} & q^{2(D-1)} & & q^{(D-1)^2} \end{pmatrix}, \quad (33)$$

where $q = \exp(2\pi i/D)$. The matrix elements of T are

$$T_{rc} = \frac{1}{\sqrt{D}} \exp\left(\frac{2\pi i}{D}(r \times c)\right), \quad (34)$$

where $r, c \in [0, \dots, D-1]$.

For $D = 2$ Jacobs' protocol [4] turns out to be equivalent to the QFT protocol. The transformation unitaries in both cases are

$$T_{\mathrm{Jac}} = \exp\left(i\frac{\pi}{2}J_y\right) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad (35)$$

$$T_{\mathrm{QFT}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad (36)$$

The effect of these transformations on dL is identical: $\check{X} = T_{\mathrm{Jac}} J_z T_{\mathrm{Jac}}^\dagger = T_{\mathrm{QFT}} J_z T_{\mathrm{QFT}}^\dagger = J_x$. Because this scheme is equivalent to Jacobs we will not analyse this $D = 2$ case any further.

For a system of dimension $D = 3$, the transformed measurement is

$$\check{X} = T J_z T^\dagger = \frac{1}{3} \begin{pmatrix} 0 & 1-q & 1-q^* \\ 1-q^* & 0 & 1-q \\ 1-q & 1-q^* & 0 \end{pmatrix}, \quad (37)$$

where $q = \exp(2\pi i/3)$. Taking the state to be $\rho = \mathrm{diag}(\lambda_0, \lambda_1, \lambda_2)$, where $\lambda_r > \lambda_c$ when $r < c$ and solving the equation for dL, using the fact that $L = \sum_{r \neq c, c \neq r} \lambda_r \lambda_c$, we find

$$dL = -8\gamma dt \sum_{r,c} |\check{X}_{r,c}|^2 \lambda_r \lambda_c$$

$$= -8\gamma dt \left(\frac{2}{3}(\lambda_0\lambda_1 + \lambda_0\lambda_2 + \lambda_1\lambda_2)\right) \quad (38)$$

$$= -8\gamma dt \frac{1}{3} L(t), \quad (39)$$

which coincides with the bound of Eq. (30). The above expression can be easily integrated to give

$$L(t) = e^{-\frac{8}{3}\gamma t} L(0), \quad (40)$$

In Fig 7 the weights $|\check{X}_{r,c}|^2$ between the $\lambda_r \lambda_c$'s are plotted. The weights in this case are all equal, so it is obvious that permuting the eigenvalues will not change the decrease of impurity. This explains why we did not have to find an optimal permutation.

Now we compare the above calculation to numerics. The average impurity for the continuous complementary
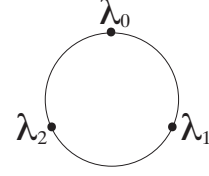

FIG. 7: The only possible permutation for $D = 3$. The solid lines between the eigenvalues represent the weighting of the multiplying factors in Eq. 38. That is, $\frac{2}{3} = |\check{X}_{r,c}|^2$, $\forall r, c$.
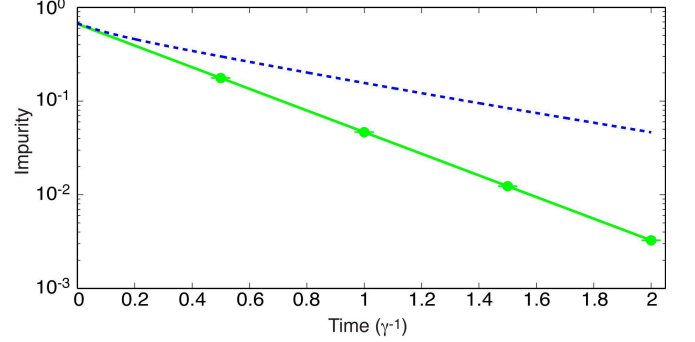

FIG. 8: Trajectory simulations of the impurity for a $D = 3$ system subject to feedback every $\delta t = 10^{-3} \gamma^{-1}$ to keep the state and measurement eigenbases complementary. The ensemble size is 100. The dashed line is $\langle L \rangle$ for a continuous commuting measurement ($D = 3$) $i.e.$ Eq. (15). The solid line is the numerically calculated ensemble average of $\langle L \rangle$ for the complementary measurement; it is plotted in green; error bars are plotted at times $(0.5, 1, 1.5, 2) \times \gamma^{-1}$. The analytic expression for the average impurity, Eq. (40), is indistinguishable from the results of the simulation.

measurement protocol and continuous commuting measurement are calculated by stochastic simulations and Eq. (15) respectively. Figure 8 shows that the stochastic simulations of the complementary measurement protocol give a deterministically decreasing $\langle L \rangle$ as predicted by Eq. (40).

We numerically investigated the effect of varying the frequency, $1/\delta t$, of the operation applied (feedback) to keep the state and measurement eigenbases complementary in Fig. 9. The stochastic fluctuations in the simulations arise from the finite size of δt in the simulations. This effect can also be found in simulations of the Jacobs' qbit feedback protocol [60]. As $\delta t \to 0$ the numerically calculated average impurity for complementary measurement approaches Eq. (40). For $\delta t \ll 10^{-3}\gamma^{-1}$ numerical simulations are indistinguishable from the analytically calculated (then interpolated) speed-up, the solid line in Fig. 9.

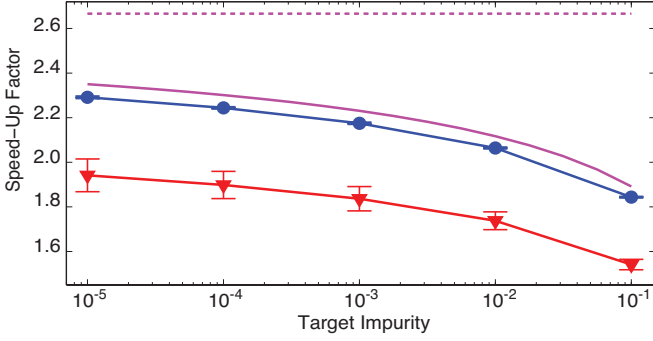We now consider complementary measurement using the QFT for the case when $D = 4$. The increment for

FIG. 9: Short time numerically calculated speed-up for $D = 3$. The dashed line is the analytically calculated asymptotic speed-up $S = 8/3 = 2.\dot{6}$. The solid (magenta) line is the ideal speed-up; calculated numerically from equations Eq. (15) and Eq. (40). The circles and triangles are numerical calculations of the speed-up with finite $\delta t = 10^{-3}\gamma^{-1}$ and $10^{-2}\gamma^{-1}$ respectively; the ensemble size is 100. As $\delta t \to 0$ the numerically calculated speed-up approaches the solid line and the error-bars vanish.

the impurity is

$$
\begin{aligned}
dL &= -8\gamma dt \sum_{r,c} |\check{X}_{r,c}|^2 \lambda_r \lambda_c \\
&= -8\gamma dt \left(\lambda_0 \lambda_1 + \tfrac{1}{2}\lambda_0 \lambda_2 + \tfrac{1}{2}\lambda_1 \lambda_3 \right. \\
&\qquad \left. + \lambda_0 \lambda_3 + \lambda_2 \lambda_3 + \lambda_1 \lambda_2 \right).
\end{aligned} \tag{41}
$$

Unfortunately, it is not possible to factor out an ex-
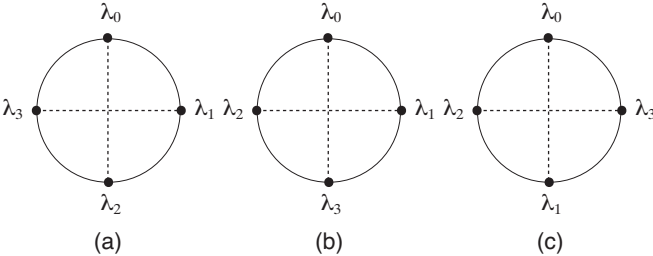

FIG. 10: (a) The weights $|X_{r,c}|^2$ in Eq. (41) after diagonalization and descending ordering. (b) The optimal permutation for $D = 4$. (c) The worst permutation for $D = 4$. The solid line represents maximum weighting of multiplication in Eq. (41), the dashed lines represent minimum weights corresponding to $\tfrac{1}{2}$.

pression for the impurity on the RHS of the above equation. It is obvious from Eq. (41) that permuting the eigenvalues will affect $|dL|$. Our task is to maximize the decrease in dL. From Eq. (41) it is possible to intuit the form of the best and worst permutations, which are depicted in Fig. 10. An optimal permutation is $\rho_{\text{opt}} = \text{diag}(\lambda_0, \lambda_1, \lambda_3, \lambda_2)$. This permutation is not uniquely optimal; the permutation $\text{diag}(\lambda_1, \lambda_0, \lambda_2, \lambda_3)$ is also optimal. One of the worst permutations is $\rho_{\text{worst}} =$

$\text{diag}(\lambda_0, \lambda_3, \lambda_1, \lambda_2)$. Unfortunately knowing the optimal permutation does not help us simplify the expression. Because we cannot solve this case analytically, we invoke the procedure developed in Refs. [5, 9] to find the bounds on $|\langle dL\rangle|$ (and hence S) for the QFT protocol for all D. These bounds are calculated in Section IV C.

Figure 11 depicts the numerically calculated speed-up for $D = 4$. The speed-up is larger than that of the $D = 3$ case [61]. Furthermore, for reasonably frequent feedback the numerically calculated speed-up lies between the analytical bounds predicted by Eq. (30) and Eq. (31).
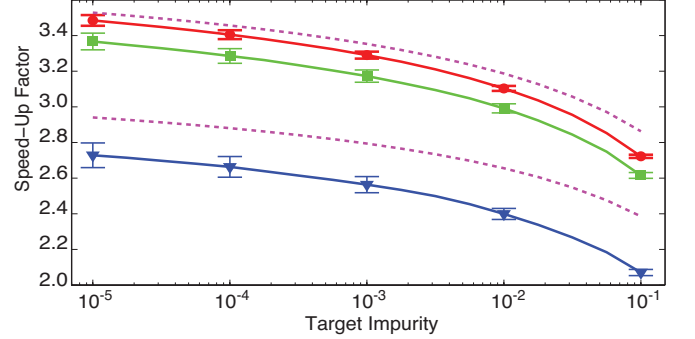

FIG. 11: Short time numerically calculated speed-up for $D = 4$. The dashed lines are the bounds on the speed-up given by Eq. (44) and Eq. (48). The circles, squares and triangles are numerical calculations of the speed-up with finite $\delta t = 10^{-3}\gamma^{-1}, 10^{-2}\gamma^{-1}, 10^{-1}\gamma^{-1}$ respectively. To keep the error bars roughly the same size the ensemble size was varied. The ensemble sizes for circles, squares and triangles are 100; 400; and 6,400 respectively.

·

We also briefly compare the complementary measurement protocol to the rapid measurement (RM) algorithm of Ref. [5] in Fig. 12. The decrease in the ensemble average of the impurity seems to be of the same order. However, the trajectories in the RM case have a large variance. This is because the trajectories are not differentiable. By comparison, even for a finite δt the L in complementary measurement is very close to *deterministic*. Although L is still stochastic, it is also differentiable for $\delta t \to dt$ so the noise in the simulations is reduced.

C. Bounds on the QFT complementary measurement protocol for all D

1. Lower bound

From Eq. (32) it is clear that the lower bound on the optimal speed-up for measurement any complementary basis is $S = \frac{2}{3}(D + 1)$. Nevertheless it is interesting to work this out explicitly for a particular complementary basis. The method we use to find the lower bound was first presented in Ref. [9]. This method uses a fictitious
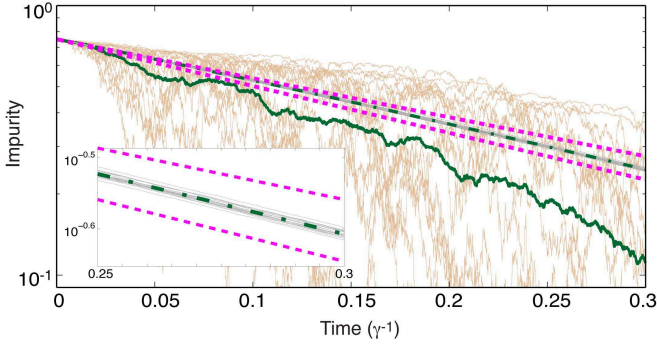
FIG. 12: A comparison between the complementary measurement protocol and rapid measurement (RM) feedback (of Ref. [5]) for $D = 4$; $\delta t = 10^{-4}\gamma^{-1}$; and an ensemble of 20 trajectories in each case. The dashed lines are the analytically calculated bounds on dL for complementary feedback [Eq. (44) and Eq. (47)]. The dot-dashed line is the ensemble average for L with QFT feedback. The solid (dark green) line is the ensemble average for the RM feedback. The light brown lines are the trajectories for the RM feedback. The analytical bounds suggest that both feedback strategies have the same order of speed-up. Inset figure: all of the complementary measurement trajectories (grey lines) lie within the predicted bounds. The trajectories are very nearly deterministic.

state ρ_F with the same purity as the state under consideration, i.e. $L[\rho] = L[\rho_F]$. The fictitious state ρ_F is defined as $\rho_F = \mathrm{diag}(1 - \Delta, \delta, .., \delta)$ where $1 - \Delta$ is the largest eigenvalue of ρ_F and $\delta = \Delta/(D-1)$. The equation for dL is invariant under permutations of the eigenvalues ρ_F. Substituting ρ_F into the RHS of Eq. (29) gives

$$2(1 - \Delta)\frac{\Delta}{(D - 1)} \sum_{r \neq 1} |\check{X}_{r1}|^2 + \frac{\Delta^2}{(D - 1)^2} \sum_{r \neq c, c \neq r} |\check{X}_{rc}|^2.$$

This expression is further simplified by explicitly calculating the matrix elements. For the QFT we have

$$\sum_{r \neq 1} |\check{X}_{r1}|^2 = \sum_{m=1}^{D-1} \frac{1}{4\sin^2(\pi m/D)} = (D^2 - 1)/12, \quad (42)$$

using the identity $\sum_{m=1}^{D-1} \csc^2(\pi m/D) = (D^2 - 1)/3$ [35]. Also,

$$\sum_{r \neq 1, c \neq r} |\check{X}_{r,c}|^2 = \frac{D + 1}{12}(D - 1)(D - 2), \quad (43)$$

using the above identity and the fact that $\csc^2(\pi(D - 1)/D) = \csc^2(\pi/D)$. After some simplification the final expression is an upper bound on the decrease in impurity:

$$dL_F \leq -\frac{2}{3}(D + 1)\gamma dt L(t), \quad (44)$$

which implies that the lower bound on the asymptotic speed-up is

$$S_F \geq \frac{2}{3}(D + 1), \quad (45)$$

as found in Refs. [8, 9] for a general complementary measurement.

2. Upper bound

To find the upper bound on the speed-up for the QFT protocol we again uses the method of Ref. [9]. This instance requires a different fictitious state ρ_2, again with $L[\rho] = L[\rho_2]$, which was termed the binary distribution. The binary distribution is defined as $\rho_2 = \mathrm{diag}(1 - \Delta', \Delta', 0, ..., 0)$ where $1 - \Delta'$ is the largest eigenvalue. It is known that dL is most sensitive to permutations of the eigenvalues of ρ_2 [9].

Substituting ρ_2 in to Eq. (29) gives only two terms. Now there is a choice about where to situate the two eigenvalues so that dL is maximized. By calculating the element $|\check{X}_{r,c}|^2$, using the QFT, we find that

$$|\check{X}_{r,c}|^2 = \frac{1}{2\left(1 - \cos\frac{2\pi}{D}(r - c)\right)}. \quad (46)$$

This allows us to optimisze over the permutations by finding $\max |P_m^\dagger \check{X}_{r,c} P_m|^2$. The largest element of the matrix is: $|\check{X}_{01}|^2 = 1/2\left(1 - \cos\frac{2\pi}{D}\right)$. From this it is clear that dL will be maximized provided the two eigenvalues from ρ_2 are in succession. The change in the impurity for the binary distribution is

$$dL_2 = \frac{-8\gamma dt(2\lambda_0\lambda_1)}{2\left(1 - \cos\frac{2\pi}{D}\right)} = \frac{-4\gamma dt}{\left(1 - \cos\frac{2\pi}{D}\right)}L_2(t). \quad (47)$$

Thus the asymptotic speed-up upper bound is

$$S_2 \leq \frac{4}{1 - \cos\frac{2\pi}{D}}. \quad (48)$$

For $D \gg 1$ one finds

$$S_2 = \frac{2D^2}{\pi^2} \approx 0.2D^2. \quad (49)$$

In Fig. 13 we depict the conjectured optimal and worst permutations of eigenvalues of a D dimensional state ρ. The optimal permutation is $\rho = \mathrm{diag}(\lambda_0, \lambda_1, \lambda_3, \lambda_5, \cdots, \lambda_6, \lambda_4, \lambda_2)$.

In Fig. 14 the numerically calculated asymptotic speed-up as a function of D is plotted. As expected, it is within the lower bound of Eq. (45) and the upper bound of Eq. (48). Further it confirms a very nearly quadratic speed-up: the fit shown is $S = 0.189D^2 + 0.109D + 0.248$. If, however, we choose to only fit the quadratic term, the fit is

$$S = 0.19D^2. \quad (50)$$

For $D \geq 10$ this is a good approximation and is very close to Eq. (49), which was derived for long-times and large D. The numerically calculated asymptotic speed-up is only a multiplicative constant away from the ultimate upper bound on all complementary basis purification protocols, the dashed line ($S = 0.5D^2$), i.e. Eq. (32).
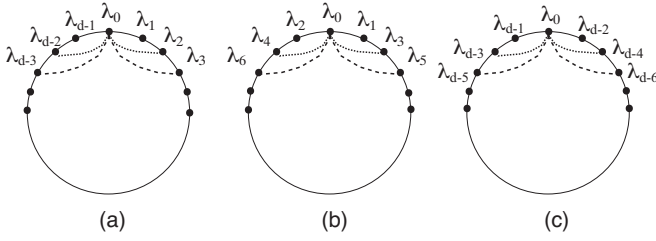
FIG. 13: a) Weighting factors in Eq. (29) after the feedback has diagonalized and ordered the eigenvalues in ρ. b) The conjectured general optimal permutation. c) The conjectured worst permutation. The weights are only shown from the λ_0 perspective.
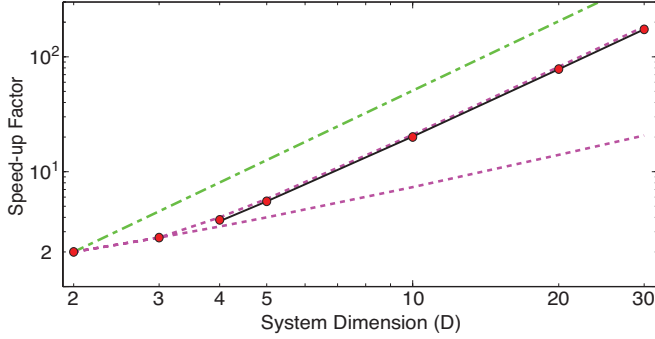

FIG. 14: The asymptotic speed-up for reaching a given level of purity as a function of the system dimension. The solid line the fit in Eq. (50). The dashed lines are the upper, Eq. (48), and lower, Eq. (45), bounds derived in the text for the QFT protocol. The dot-dashed line is the upper bound on all unbiased basis purification strategies, that is the RHS of Eq. (32).

D. Discussion

The results presented in sections Sec. IV A–Sec. IV C raise a number of questions which we now elucidate further.

Because of symmetries, for dimensions two and three permuting the basis of ρ was unnecessary to maximize dL. The intuition is, for maximizing information gain the complementary measurement marginalises the significance of which eigenvalue is the largest. For $D > 3$ we must permute the basis of ρ, after diagonalization and before applying T_{QFT}, to maximize dL. The permutations applied do not increase the signal to noise ratio of the measurement unlike the rapid measurement protocol of Ref. [5]. In Sec. IV B and Sec. IV C 2 we argued that the permutations should be chosen to maximize the product of the two largest eigenvalues in Eq. (29). While this recipe is operationally sound it lacks physical insight. For example, this recipe does not explain what the mechanism of the purification is and why the permutations can dramatically affect the rate of purification. These questions are addressed in Sec. V.

The final open question concerns the efficacy of different complementary bases to change the rate of purification. In previous work Lund and Jacobs suggested that not all complementary bases were equally good at entropy reduction [19]. That is, they claimed that one particular unbiased basis can reduce entropy more quickly than another. If this were true then the $S = 0.2D^2$ upper bound we derived from the QFT basis would not necessarily be the true upper bound, which could be closer to $S = 0.5D^2$ We numerically investigate this question in Sec. VI.

V. WHY COMPLEMENTARITY IS NOT ENOUGH

How is it that rotating to an unbiased basis provides any speed-up? Naively following the reasoning applied to qbits in Refs. [2–4] suggests that making each outcome equally likely (*i.e.* maximizing the surprisal [1]) one maximizes the average amount of information a measurement extracts. From this one might argue that the speed-up observed can be explained by the D-dimensional version of this argument. However, it is not clear how this argument would explain why the arrangement of eigenvalues in ρ (permutations) are important to attain the best speed-up. To address this, we examine a phase space representation of the optimal and worst permutations that are schematically depicted in Fig. 13.

Defining a phase space picture for a discrete variable is not trivial. In Ref. [36] it was shown that a spin Wigner function $W(\theta, \phi)$ can be defined in terms of Clebsch-Gordan coefficients and spherical harmonic functions. This spin Wigner function is a pseudo-probability distribution on the Bloch sphere, with θ and ϕ the usual Euler angles. The spin Wigner function is a little counter-intuitive; for example, unlike the original Wigner function, $W(x, p)$, for position and momentum the marginal distribution for ϕ is not the true phase distribution $P(\phi)$. However, for large D the marginals are a good approximation to the true phase distribution [37]. We plot the Wigner function using the equal-area projection (described by co-ordinates ϕ and $J\cos\theta$).

The conjectured best and worst permutations were depicted on a ring in Fig. 13, explained in terms of the weighting factors $|X_{r,c}|^2$. We now have an intuitive understanding of the angle around this ring as being the phase ϕ in the angular representation of spin states. The eigenstates of ρ in the measurement (J_z) basis are $|r\rangle$ (Dicke states [38]). These are transformed by the QFT to the states $|\phi_r\rangle$, where $\phi_r := \frac{2\pi}{D}(J - r)$. The states $|\phi_r\rangle = \frac{1}{\sqrt{D}} \sum_{m=-J}^{J} \exp(-im\phi_r)|J, m\rangle$ are equivalent to the Pegg-Barnett phase states [39, 40]. If the QFT were an easy operation in some physical system then our protocol would be a procedure for rapidly preparing a Pegg-Barnett phase state. We note that in some spin systems it should be possible to construct the desired unitary [41].

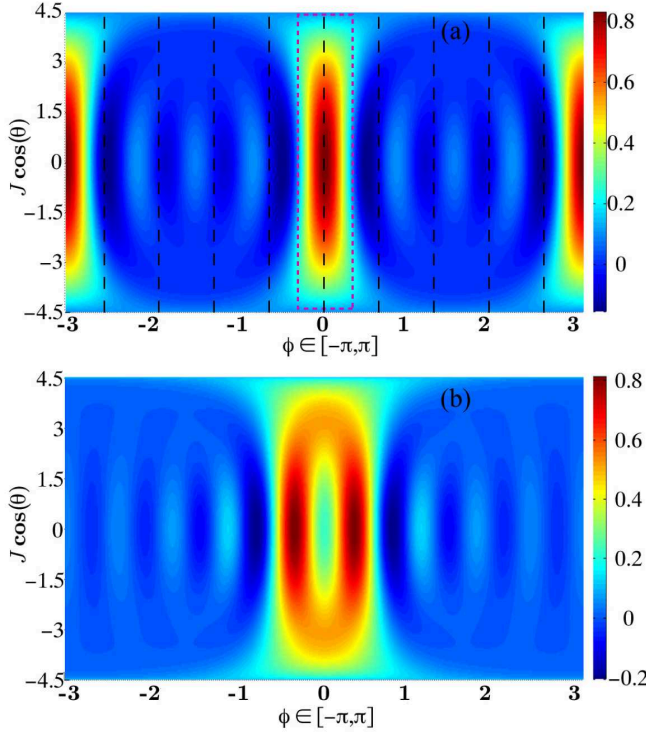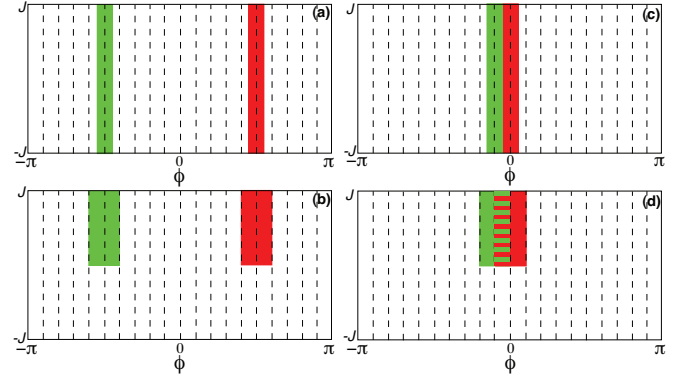Consider the long-time-limit state. The worst case, for

FIG. 16: Schematic Wigner functions for mixtures of two phase eigenstates for $D = 20$. (a) The worst permutation cf. Fig. 13 (c); we have rotated the entire sphere by $\pi/2$ about z to simplify the explanation of the speed-up. b) The worst permutation after a positive measurement result (modulo $\pi/2$). The total area is constant in figures (a) and (b), and is equal to 4π. (c) The optimal permutation cf. Fig. 13 (b). (d) The optimal permutation after a positive measurement result. The total area is reduced in figure (d) to 3π, which gives rise to a purifying effect (the purity is inversely proportional the area underneath the Wigner function). The striped region $\phi \in (-\pi/D, 0]$ denotes that the red and green rectangles are overlapping in this region. From these figures it is apparent that the circles in Fig. 13 can be thought of as showing ϕ.

FIG. 15: An equal area plot of the Wigner function of an equal mixture of two eigenstates after the QFT, $D = 10$. a) the worst permutation cf. Fig. 13 (c). The dashed vertical lines correspond to the locations of the peaks of the D transformed eigenstates. The dotted box around the peak at $\phi = 0$ corresponds to the bulk of the Wigner function for this state. b) The conjectured general optimal permutation, cf. Fig. 13 (b).

purification purposes, is when the eigenvalues are equal. Under this two eigenvalue approximation the worst permutation corresponds to putting the second largest eigenvalue, λ_1, the maximal distance away from λ_0 in phase space ($\phi = \pm\pi$). The spin Wigner function for this configuration of the mixture is plotted in Fig. 15 (a) for $D = 10$. The optimal permutation is when the two largest eigenvalues are next to each other in phase space; see Fig. 15 (b). It is now apparent that the schematic diagrams in Fig. 13 represent a slice through the unwrapped Bloch sphere and the positions of the eigenvalues are their arrangements of the phase states corresponding to the original eigenvalues. This phase space picture suggests it might be possible to explain why these are the best and worst cases and how the speed-up is generated. To explain these two features we move away from exact calculation of Wigner functions and move to a schematic representation of the Wigner function.

In Fig. 16 (a) we represent schematically the bulk of the Wigner function by a rectangle of width $2\pi/D$ (corresponding to the ϕ coordinate) and height $D - 1 = 2J$ (which corresponds to the $J\cos\theta$ coordinate). In what follows one may loosely think of the rectangle as rep-

resenting the uncertainties in an observer's knowledge about two conjugate variables; according to the Heisenberg uncertainty relation (HUR) the area of the rectangle must be constant for pure states. Now consider the effect of a weak measurement of J_z. In our protocol this is of infinitesimal duration, but here we exaggerate the effect to illustrate our point. In Fig. 16 (b) we have taken the result of the measurement to be positive. Because the result was positive, the Wigner function does not have much support on the lower part of the plot. As the positive result contained information about the distribution of j, the uncertainty in this variable is reduced. In keeping with the HUR, the conjugate variable (ϕ) suffers an increase in variance. For Wigner functions the purity is proportional to the inverse of area under the function [42]. Here there has been no change in the total area and hence no change in the total purity. Because this is only a heuristic for understanding protocol the previous statement is not entirely true; the final paragraph of this section will explain the actual result.

In Fig. 16 (c) the optimal permutation in the QFT basis is depicted. The total area of the two rectangles before the measurement is $\sim 4\pi$. After a positive measurement result the phase distributions for the two eigenvalues significantly overlaps. The total area is now $\sim 3\pi$; this reduction in area leads to an increase in purity; see Fig. 16 (d). Now the intuitive understanding for the speed-up and the permutation sensitivity is apparent. The per-

mutations are important so that the large eigenvalues in ρ may bleed into each other after a measurement. The bleeding is due to measurement backaction in the variable conjugate to J. It is the reduction in area this bleeding effects that causes the purification; but it only works if the largest two eigenvalues are adjacent in phase space. Although this picture is crude it captures the essence of the protocol.

So far we have given an intuitive explanation of the mechanism of the purification. Now we provide an intuitive explanation for the speed-up. Consider the integrated measurement result

$$\int_0^t dR(t) = \int_0^t \sqrt{4\gamma} X(s) ds + \int_0^t dW(s), \qquad (51)$$

this expression contains a term corresponding to the signal of interest (the first term) and a term representing the noise (the second term). We may define the signal to noise ratio as $\text{SNR} = \text{signal}^2/\text{noise}^2$. In this case we have $\text{signal}^2 = 4\gamma \Delta t^2 \langle X^2 \rangle$ and $\text{noise}^2 = \langle (\Delta W(t))^2 \rangle = \Delta t$. Thus the ratio becomes $\text{SNR} = 4\gamma \Delta t \langle X^2 \rangle$. Now we wish to estimate the time taken to evolve from the Wigner functions in the first row of Fig. 16 to those in the second row. Note that an observer gains one bit of information about J_z between the first and second rows. From the Shannon–Hartley theorem, $\text{bits} = \log_2(1 + S/N)$, we infer that this implies $\text{SNR} = 1$. Now we solve for time and take $X = J_z$ consequently $\langle J_z^2 \rangle = (D^2 - 1)/12$ (see appendix C in Ref. [9]) to find

$$\Delta t = \frac{3}{\gamma(D^2 - 1)}. \qquad (52)$$

Recall that the purity is approximately inversely proportional to the phase space area. The initial purity in Fig. 16 (c) is $P_i \propto 1/4\pi$ and we know the state is an equal mixture of two eigenstates which means the purity equals one half. This allows us to determine the proportionality constant to be 2π, so $P_i = \frac{1}{2}$. The final purity is thus $P_f = 2/3$ which means the change in purity is $\Delta P = 1/6$. Consequently

$$\frac{1}{P}\frac{\Delta P}{\gamma \Delta t} = \frac{1}{2}\frac{(1/6)}{3/(D^2 - 1)} = \frac{1}{9}(D^2 - 1), \qquad (53)$$

This can be compared to Eq. (49) that is

$$\frac{1}{L}\frac{dL}{\gamma dt} = -\frac{2D^2}{\pi^2} \approx -\frac{1}{5}D^2. \qquad (54)$$

Considering the crudeness of the arguments we have employed, there is surprisingly good agreement between Eq. (53) and Eq. (54). This gives weight to the intuition that the effect of purification for the QFT protocol comes from measurement back-action and the role of the permutations is to maximize the back-action by placing the eigenvalues close in phase space.

We now return to the worst permutations for purification. All one can conclude from the equi-likelihood

of all outcomes is that there should be a speed-up of at least two (in a discrete outcome measurement model, which is equivalent to our current measurement model). Let us confirm this intuition now. Returning to the two eigenvalue approximation, consider the worst permutation. It is possible to determine the upper bound on speed-up resulting from this permutation by substituting $r - c = D/2$ into Eq. (46). We find that $S = 2$, which confirms the intuition above. This is why complementarity is not enough to guarantee maximal information extraction.

VI. ARE ALL COMPLEMENTARY BASES EQUALLY GOOD FOR ENTROPY REDUCTION?

We now address the question raised by Jacobs and Lund [19]: are all unbiased bases equally good for rapid purification? Their answer was no. However at the time of their analysis the role of permutations was not clearly understood. Thus it is worthwhile to re-examine this question. Due to rotational symmetry of the unbiased bases and the permutational symmetry of density operator the answer in the cases $D = 2$ and 3 is yes. However for $D = 4$ it is easy to find a counter-example to this trend. For example measuring in any of the four *mutually unbiased bases* (MUBS) [43] gives

$$
\begin{aligned}
dL &= -8\gamma dt \sum_{r,c} |\breve{X}_{r,c}|^2 \lambda_r \lambda_c \\
&= -8\gamma dt \left[2\lambda_0 \lambda_1 + \tfrac{1}{2}(\lambda_1 \lambda_2 + \lambda_0 \lambda_3) + 2\lambda_2 \lambda_3 \right. \\
&\qquad \left. + 0(\lambda_0 \lambda_2 + \lambda_1 \lambda_3) \right],
\end{aligned}
\qquad (55)
$$

here $\breve{X} = M_i^\dagger J_z M_i$ where M_i is a transformation to one of the MUBS. (Explicit expressions of the five MUBS when $D = 4$ are given in Ref. [43] for example. For convenience we have reproduced them in a footnote [62].) As before, we may use ρ_2 to obtain a lower bound on dL (and hence an upper bound on the speed-up). Doing so gives $dL = -8\gamma dt L(t)$ which implies a speed-up of $S = 8$. This saturates the upper bound on the speed-up found in Eq. (32). The factor of two improvement over the QFT can be understood from the difference in weights of the largest terms $\lambda_0 \lambda_1$ between Eq. (41) and Eq. (55).

To get more intuition about the purification process using the MUBs transform we look at the Wigner function of the four states in one of the four MUBs, which for convenience we denote $\{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}$, in Fig. 17. The states plotted in Fig. 17 (a) and (d) look like the states in the QFT basis (Pegg-Barnett phase states). The Wigner functions plotted in Fig. 17 (b) and (c) are however quite different — they contain a hole which is more negative than the dips in Fig. 17 (a) and (d). The states plotted in Fig. 17 (b) and (c) can be said to be highly non classical because of this.

We now compare the MUB transform to the QFT using the Wigner function, in an attempt to get an intuitive explanation of the advantage of the MUB transform.
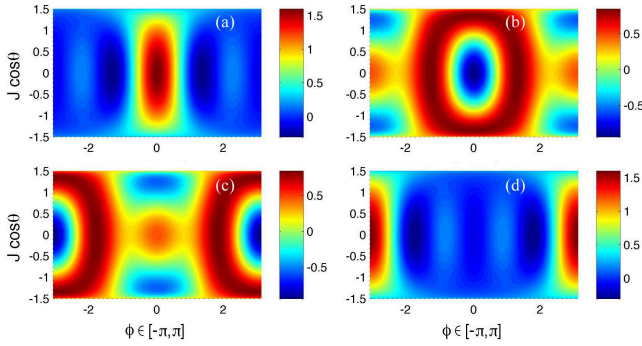
FIG. 17: An equal area plots of the Wigner function resulting from transforming four different states to the M_1 unbiased basis ($D = 4$). (a) the transform of $\rho_0 = \mathrm{diag}(1, 0, 0, 0)$ (b) $\rho_0 = \mathrm{diag}(0, 1, 0, 0)$ (c) $\rho_0 = \mathrm{diag}(0, 0, 1, 0)$ and (d) $\rho_0 = \mathrm{diag}(0, 0, 0, 1)$. Unlike the unbiased basis generated by the QFT not all of the resulting states are like the Pegg-Barnett phase states; see for example figures (b) and (c).

FIG. 18: A Wigner function comparison between the QFT (top row) and MUB transform (bottom row) for $D = 4$. We plot an equal mixture of two eigenstates with the optimal permutation (first column) and a poor permutation (second column) for each transform. In figures (a)–(d) the black dashed line is the zero contour. (a) the optimally permuted state for the QFT – $\rho = \mathrm{diag}(1, 1, 0, 0)/2$; (b) the worst permutation for the QFT – $\rho = \mathrm{diag}(1, 0, 1, 0)/2$; (c) the optimally permuted state for the MUB transform – $\rho = \mathrm{diag}(1, 1, 0, 0)/2$; (d) a poor permutation for the MUB transform – $\rho = \mathrm{diag}(1, 0, 0, 1)/2$. Note that cases (b) and (d) are actually identical. Unlike the QFT case, the mixture of optimally permuted MUB states has two peaks which could be distinguished by a measurement of J_z. (e) The worst permutation *i.e.* $\rho = \mathrm{diag}(1, 0, 1, 0)/2$ after the transform to the M_1 unbiased basis, which results in *no* purification.

To make the analysis simple we restrict to states which have two large eigenvalues only (as we did above). From Eq. (55) we see that permuting these states in the logical basis so that they correspond to $|0\rangle$ and $|1\rangle$ (or $|2\rangle$ and $|3\rangle$) is the optimal thing to do. Recall that permuting the two largest eigenvalues to $|0\rangle$ and $|1\rangle$ was also optimal for the QFT protocol; see Fig. 18 (a). There we found that doing so created two peaks close in phase space, so the purification could be explained by measurement backaction. In Fig. 18 (c) the Wigner function of the optimally permuted state for the MUB transform is depicted [it is an equal mixture of Fig. 17 (a) and (b)]. Here we can no longer attribute the purification mechanism to backaction. In fact it seems as though the purification effect comes from distinguishing the two peaks in the J_z distribution. The worst QFT permutation also corresponds to a poor MUB transform as evidenced in Fig. 18 (b) and (d).

Interestingly there is a permutation that is worse for the MUB transform than the permutation depicted in Fig. 18 (d). It is $\rho = \mathrm{diag}(1, 0, 1, 0)/2$; the transformed state of this permutation is depicted in Fig. 18 (e). This permutation results in no purification at all! We can see this by substituting $\lambda_0 = \frac{1}{2}$ and $\lambda_2 = \frac{1}{2}$ into Eq. (55) to get $dL = 0$. This complete lack of purification may have application in state stabilisation of states with arbitrary purity [19].

We would like to know if the upperbound on the speed-up is saturated in all dimensions. To answer this question we resort to a numerical search for $D \in [2, 10]$. We do not claim that our search is exhaustive. Recall that the stochastic simulations show that speed-up predicted by ρ_2 is close to the achievable amount. It seems reasonable to assume this to be true in other complementary bases. This greatly simplifies the analysis by making the optimisation of permutations superfluous. Because of this
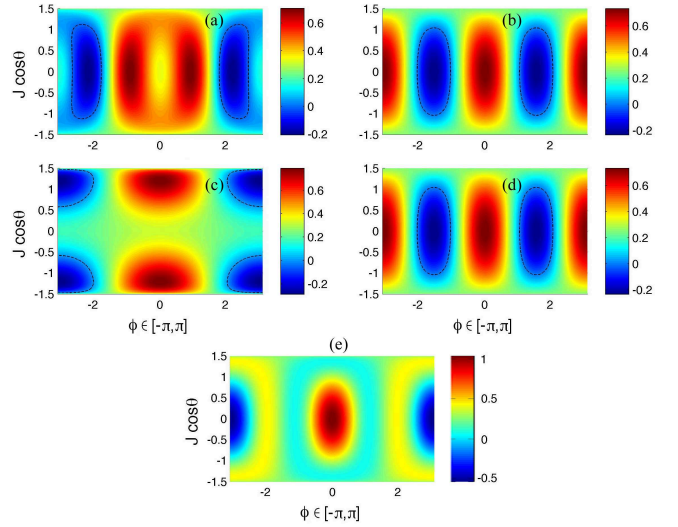
we may numerically search for over all unbiased bases for the basis which has the largest element $|\check{X}_{r,c}|^2$. Converting the element $|\check{X}_{r,c}|^2$ to a speed-up gives the following trend. The speed-up for even D attains the upperbound, *i.e.* $D^2/2$, and the speed-up for odd D equals $(D-1)^2/2$, as seen in Fig. 19. In either case this is much larger than the speed-up of the QFT protocol which was $S = 0.2D^2$. For $D \gg 1$ it is reasonable to believe that the achievable speed-up does indeed scale like the predicted $S = 0.5D^2$ [9].

VII. INFORMATION ACQUISITION IN A REGISTER OF QBITS AND ITS RELATIONSHIP TO COMPLEMENTARITY

Finally, we consider complementary measurements of a register of n qbits, where each qbit is independently and weakly (or continuously) measured, as introduced in Ref. [5]. Instead of one observable X, there are now n, given by $X^{(r)} = I^{(1)} \otimes I^{(2)} \otimes \ldots \sigma_z^{(r)} \ldots \otimes I^{(n)}$, where r labels the rth qbit. The SME describing such a measure-
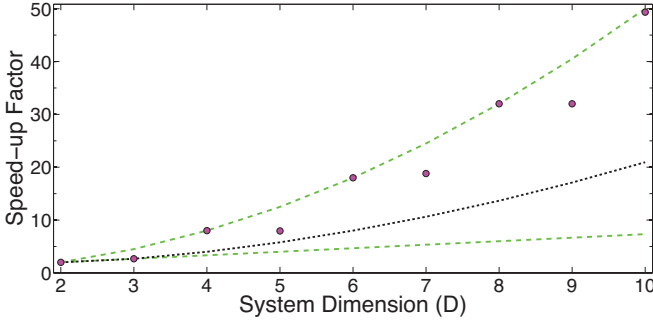
FIG. 19: The asymptotic speed-up for reaching a given level of purity as a function of the system dimension. The dashed lines are the upper bound and lower bounds on all unbiased basis purification strategies, *i.e.* Eq. (32). The dotted line is the maximum achievable speed-up for the QFT protocol *i.e.* Eq. (48). The solid circles are the results of a numerical search over unbiased bases to maximize the asympototic speed-up. See the text for more details.

ment is

$$d\rho = \sum_r 2\kappa \, dt \, \mathcal{D}\left[X^{(r)}\right]\rho + \sqrt{2\kappa}\, dW^{(r)}\mathcal{H}\left[X^{(r)}\right]\rho. \quad (56)$$

The combined state of the n qbits exists in a $D = 2^n$-dimensional Hilbert space.

A. Commuting measurements analysis

Here we will not analyse the commuting measurements of a register of qbits with the same detail as we did for the qdit in Sec. III. Instead we rely on a result from Ref. [9] where an expression for the evolution of the average impurity of a register by undergoing a continuous commuting measurement was found. In the long time (LT) limit the impurity is

$$\langle L^{(n)}(t)\rangle_{\mathrm{LT}} = \frac{n\pi e^{-4\kappa t}}{8\sqrt{\pi \kappa t}}. \quad (57)$$

We will use this expression in this section to calculate the speed-up, so only the asymptotic scaling is important: $\langle L^{(n)}(t)\rangle \sim e^{-4\kappa t}$.

B. Complementary measurements

It was shown in Refs. [9] that the change in impurity for a register of qbits monitored in an unbiased basis is

$$dL = -8\kappa \, dt \sum_{r=1}^{n} \mathrm{Tr}\left[\check{X}^{(r)}\rho\check{X}^{(r)}\rho\right] \quad (58)$$

$$= -8\kappa \, dt \sum_{r=1}^{n} \sum_{i,j=0}^{(D-1)} |\check{X}_{i,j}^{(r)}|^2 \lambda_i \lambda_j. \quad (59)$$

Here we define the complementary observable to be $\check{X}^{(r,m)} = P_m T X^{(r)} T^\dagger P_m^\dagger$. As before, the T's are conditional unitaries that introduce the unbiasedness (between ρ and $X^{(r)}$), and the P_m's are the permutation operators.

It was found that the upper bound on the impurity was

$$L_{\mathrm{UB}}^{(n)}(t) = \exp\left(-8\kappa n t/[D-1]\right)L(0), \quad (60)$$

while the lower bound was

$$L_{\mathrm{LB}}^{(n)}(t) = \exp\left(-8\kappa n t\right)L(0). \quad (61)$$

From these equations we can infer the following bounds on the asymptotic speed-up factor

$$\frac{2n}{D-1} \leq S \leq 2n. \quad (62)$$

For all $n \geq 3$ the lower bound on the speed-up becomes less than unity. For large n the slow-down implicit in the lower bound is $\sim n2^{-n+1}$. (The interested reader should also see the related study of Hill and Ralph [44].)

C. Complementarity via the quantum Fourier transform

From Eq. (62) it is not clear if measurements in a complementary basis provide any benefit in information extraction. In this subsection we present some progress towards answering this question.

We begin our analysis by considering the speed-up when $n = 2$. The bounds on the speed-up are $4/3 \leq S \leq 4$. In order to see if the upper bound is achievable we perform stochastic simulations. In Fig. 20 we numerically determine the advantage of the QFT feedback protocol in a register of two qbits. The permutation used for simulating the evolution was $\rho = (\lambda_0, \lambda_4, \lambda_3, \lambda_1)$.

In Fig. 20 (a) the QFT protocol clearly does not saturate the bound on $\langle L\rangle$ given in Eq. (61). One may randomly permute the eigenvalue arrangement at times δt before applying the QFT feedback to deterministically achieve the lower bound of Eq. (62). From Fig. 20 (b) it seems as though the speed-up is asymptoting towards 2 rather than the upper bound specified by Eq. (62), i.e. 4. For comparison, we note that the speed-up found for the locally optimal rapid measurement protocol in Ref. [5] was $S_{\mathrm{RM}} \approx 1.4$ in the long-time limit.

The structure of unbiased bases for a register of qbits is quite complicated [45–47]. Even for the QFT the optimal permutation is not obvious for $n > 2$. It is for this reason we plot the two eigenvalue approximation to the asymptotic speed-up for the QFT in Fig. 21 as a function of n. The values were obtained by finding the largest term in Eq. (59). This term will be denoted by $X_{\mathrm{max}} := \max_{i,j} \sum_r |X_{i,j}^{(r)}|$. The feedback places the two eigenvalues at i_{max} and j_{max}. Thus X_{max} is proportional to the speed-up. Curiously the speed-up,
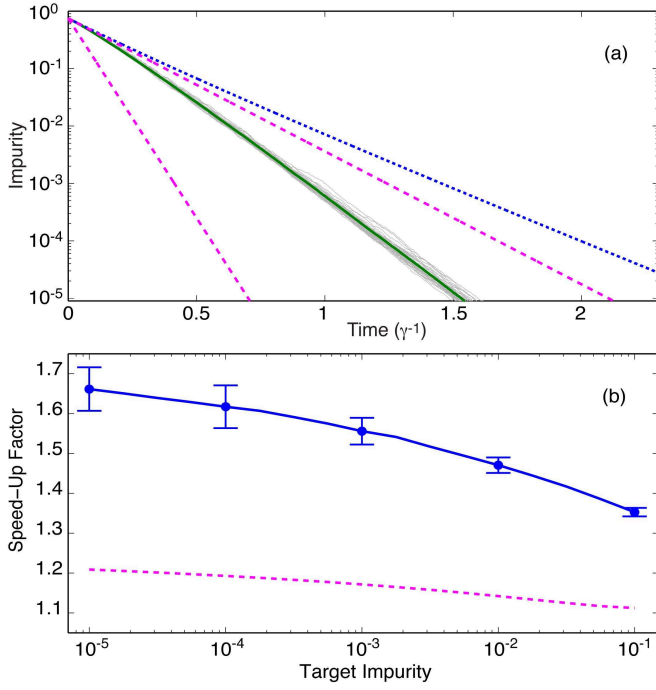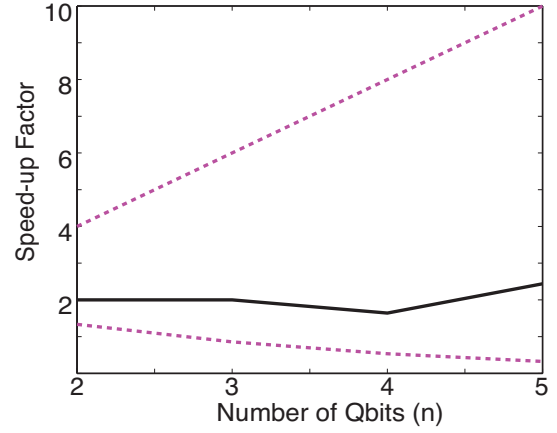
$$S \approx 2, \quad (63)$$

FIG. 21: Bounds on the asymptotic speed-up for a register of qbits. The dashed lines are the upper and lower bounds derived in the main text Eq. (62). The solid line is the largest element of $X_{\max} = \max_{i,j} \sum_r |X_{i,j}^{(r)}|$. The element $X_{\max}$ is conjectured to be a good indicator for maximum achievable speed-up with the QFT in a register of qbits.

FIG. 20: (a) Impurity for a two qbit register: dotted line – a continuous measurement; solid green line – average impurity for feedback; solid grey lines – individual trajectories grey lines; dashed magenta lines - upper [Eq. (60)] and lower [Eq. (61)] bounds. (b) Speed-up solid line; lower bound dashed line; upper bound not shown as it well above the numerical result. The ensemble size is 20 and feedback was applied at times $\delta t = 1 \times 10^{-3} \gamma^{-1}$.

appears to be independent of the size of the register. It is not clear if this is true for any unbiased basis in a register. Our arguments from Sec. VI indicate it might be a multiplicative constant higher.

VIII. DISCUSSION

Prior to this work it has been shown that it is possible to speed-up the extraction of information from a quantum system using a continuous complementary measurement [4, 8, 9]. In this paper we have given an explicit method for constructing such a method, using the QFT, that achieves a speed-up of $S = 0.2D^2$ for a qdit. This explicit construction allowed us to show that complementary between the state and measurement observable is not enough to guarantee maximal information extraction. Choosing the right permutation of eigenstates, before Fourier transforming the state – see Fig. 13, is needed to guarantee maximal information extraction, which is contrary to what one might expect from classical information theory [1]. The effect of the permutation can change the speed-up in information extraction from its minimal value 2 to $0.2D^2$. We also argued that in the

case of the qdit, the maximum possible speed-up, predicted in Ref. [9], $S \sim D^2/2$ should be achievable. One interesting upshot of our investigation (see Sec. VI) is an explicit example of using an unbiased basis for state stabilisation [19].

In this paper we demonstrated that measurement back-action is fundamental to the purification process for the QFT protocol. We did not, unfortunately, find an information theoretic explanation of the measurement process as was found in Ref. [48]. It is not clear, at present, that the bounds presented in Sec. IV C could be obtained from more elementary reasoning about uncertainty relations [49].

In related work, Shabani and Jacobs [50, 51] found the globally optimal (in time) protocol for $D = 3$ and the locally optimal protocol for all dimensions for reduction of a quantity related to impurity. Their calculated bounds on the speed-up were $2(D-1) \leq S_{\mathrm{SJ}} \leq 2(D-1)^2$. Naively comparing their bounds to ours suggest that an unbiassed protocol is not optimal. However it is not yet clear if their upper bound is achievable for $D > 4$.

It is possible to perform a simple calculation, independent of the one presented in Ref. [50], for the absolute upper bound on the speed-up. First we calculate a lower bound on $\mathrm{E}[dL]$:

$$\mathrm{E}[dL] = -8\gamma dt \Big\{ \mathrm{Tr}\left[\rho \check{X} \rho \check{X}\right] - 2\mathrm{Tr}\left[\check{X}\rho\right] \mathrm{Tr}\left[\check{X}\rho^2\right]$$
$$+ \mathrm{Tr}\left[\rho \check{X}\right]^2 \mathrm{Tr}\left[\rho^2\right] \Big\}, \qquad (64)$$

by using the trace inequality $\mathrm{Tr}\left[ABAB\right] \leq \mathrm{Tr}\left[A^2 B^2\right]$, for A, B Hermitan. The breve on X denotes the unitary transformation $\check{X} = U^\dagger X U$, where U is *any* unitary. Note that all but the first term, $\mathrm{Tr}\left[X\rho X\rho\right]$, are are

$O(1 - L^2)$ [or $O(L^2)$]. Consequently the first term dominates E[dL] as it is $O(L)$. Next we make a two eigenvalue approximation to ρ. We find

$$\text{E}[dL_2] = -8\gamma dt \text{Tr} \left[\rho_2 \check{X} \rho_2 \check{X} \right] \tag{65}$$

$$\leq -8\gamma dt \text{Tr} \left[X^2 U \rho_2^2 U^\dagger \right] \leq 2(D-1)^2. \tag{66}$$

This implies an upper bound on the speed-up of $S = 2(D-1)^2$, as found by Shabani and Jacobs [50]. The results of a numerical search, like the one performed in Fig. 19, indicates that this bound is indeed achievable. Furthermore we can confidently say that the speed-up of the time-optimal control strategy for impurity reduction is bounded above by $S = 2(D-1)^2$. This is justified by making a two eigenvalue approximation to ρ_2 and then using the proof that Jacobs' protocol [4] is optimal [11, 50–52].

It is interesting that the rapid measurement protocol (considered in Refs. [5, 7]) and the complementary measurement protocols both afford *at most* a speed-up $O(D^2)$. They work in very different ways. The rapid measurement protocol is essentially classical in nature. It uses operations that ensure the state and measurement commute at all times. In particular the operations are permutations of the state in the measurement basis. The phase space picture for the rapid measurement protocol, in the long time limit (i.e. under the two eigenvalue approximation), would be two J_z eigenstates (Dicke states) placed at $\pm J$. As the measurements are in the z basis, the mechanism for the rapid measurement speed-up is an increase in signal to noise ratio in the measurement record. In this article we have shown that the complementary measurement protocol is essentially quantum mechanical in nature. The mechanism for the purification and speed-up is measurement backaction (at least for the QFT protocol). The two approaches also have different advantages. Rapid measurement has the advantage that it enables one to obtain information about the initial state. The complementary protocol presented here has the advantage that it provides a nearly deterministic improvement in information gathering of the current system state. In future work we will explore if it is possible to obtain information about the initial state for a continuous complementary measurement.

Finally we will speak to the practicality of implementing QFT based protocols. In recent work we showed how to analyse imperfections in a purification protocol using a feedback master equation [53] so we will not discuss such issues further here. Instead we will focus on the implementation the QFT. For qbits the QFT is a fundamental quantum logic gate known as the Hadamard gate [33], and is typically easy to implement.

For $D > 2$ implementing the QFT is more difficult, as it involves multiple logic gates (for n-qbit systems) [54] or multi-level coherent operations (for a qudit). Earlier we mentioned that in some atomic spin systems it should be possible to construct the desired QFT unitary [41]. We think, however, that solid state systems are the most likely candidate for which it would be useful to implement the ideas presented in this article. This is because in solid state systems the measurement strength γ is typically much smaller (by an order of magnitude or more) than the maximum control strength $|\alpha|$ (e.g. in Ref. [55] $1/\gamma$ is on the order of microseconds and $1/|\alpha|$ is of order nanoseconds). Consequently it is possible to imagine applying feedback continuously through the measurement process so that the eigenbasis of the state and observable are QFT pairs. Alternatively, if the maximum control strength is large enough, the feedback could be applied impulsively at discrete times, which can gives results surprisingly close to the continuous version (as noted in Fig. 11 and Ref. [7]). Further the exciting results in the field of superconducting qdits ($D \in [3, 5]$) where measurement and full unitary control has been demonstrated for a number of systems [56, 57] lead us to speculate that our proposals could be experimentally tested within the next decade.

[1] M. Tribus, *Thermostatics and Thermodynamics: An Introduction to Energy, Information and States of Matter, with Engineering Applications*, (Van Nostrand, 1961), p. 64.

[2] C. A. Fuchs and K. Jacobs, Phys. Rev. A **63**, 062305 (2001).

[3] A. C. Doherty and K. Jacobs and G. Jungman, Phys. Rev. A **63**, 062306 (2001).

[4] K. Jacobs, Phys. Rev. A **67**, 030301(R) (2003).

[5] J. Combes, H. M. Wiseman and K. Jacobs, Phys. Rev. Lett. **100**, 160503 (2008).

[6] H. M. Wiseman and J. F. Ralph, New J. Phys. **8**, 90 (2006).

[7] J. Combes, H. M. Wiseman, A. J. Scott, Phys. Rev. A **81**, 020301(R) (2010).

[8] J. Combes and K. Jacobs, Phys. Rev. Lett. **96**, 010504 (2006).

[9] J. Combes, H. M. Wiseman, K. Jacobs, and A. J. O'Connor, Phys. Rev. A. **82**, 022307(2010).

[10] C. M. Caves, C. A. Fuchs, and R. Schack, J. Math. Phys. **43**, 4537 (2002); R. W. Spekkens, Phys. Rev. A **75**, 032110 (2007).

[11] H. M. Wiseman and L. Bouten, Quantum Information Processing 7, 71 (2008).

[12] H. J. Carmichael, *An Open Systems Approach to Quantum Optics (Lecture Notes in Physics vol 18)*, (Springer Berlin, 1993).

[13] N. Gisin, Phys. Rev. Lett. **52**, 1657 (1984); C. M. Caves and G. J. Milburn, Phys. Rev. A **36**, 5543 (1987); L. Diosi, Physics Letters A **129**, 419 (1988); V. P. Belavkin,*Nondemolition measurement and nonlinear filtering of quantum stochastic processes*, (Springer, Berlin, 1988); J. Dalibard, Y. Castin, and K. Mølmer, Phys. Rev. Lett. **68**, 580 (1992); A. Barchielli and M. Gregoratti, *Quantum Trajectories and Measurements in Continuous Time: The Diffusive Case*, (Springer, Berlin, 2009).

[14] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control*, (Cambridge University Press, Cambridge 2010).

[15] W. H. Zurek, *Decoherence and the Transition from Quantum to Classical–Revisited*, Progress in Mathematical Physics **48**, 1 (2007).

[16] J. K. Stockton, R. van Handel, and H. Mabuchi, Phys. Rev. A **70**, 022106 (2004).

[17] S. Dolinar, Tech. Rep. **111**, Research Laboratory of Electronics, MIT Technical Report No. 111, (1973); H. M. Wiseman, Phys. Rev. Lett. **75**, 4587 (1995); H. M. Wiseman and R. B. Killip, Phys. Rev. A **57**, 2169 (1998); B. L. Higgins *et al.*, Nature **450**, 393 (2007).

[18] V. P. Belavkin, Rep. Math. Phys., **43**, 405 (1999); H. M. Wiseman, Phys. Rev. A **49**, 2133 (1994); A.N. Korotkov, Phys. Rev. B **63**, 115403 (2001); H. M. Wiseman and A. C. Doherty, Phys. Rev. Lett. **94**, 070405 (2005).

[19] K. Jacobs and A. P. Lund, Phys. Rev. Lett. **99**, 020501 (2007).

[20] H. M. Wiseman and S. Mancini and J. Wang, Phys. Rev. A **66**, 013807 (2002).

[21] R. Ruskov, A. Korotkov, and K. Mølmer, Phys. Rev. Lett. **105**, 100506 (2010).

[22] K. Jacobs, Proc. SPIE Int. Soc. Opt. Eng. **5468**, 355 (2004).

[23] H. M. Wiseman and G. J. Milburn, Phys. Rev. A **47**, 642 (1993).

[24] K. Jacobs and D. Steck, Contemp. Phys. **47**, 279 (2006).

[25] T. A. Brun, Am. J. Phys. **70**, 719 (2002).

[26] K. Jacobs, *Stochastic Processes for Physicists* (CUP, Cambridge, 2010).

[27] R. van Handel, J. K. Stockton and H. Mabuchi, IEEE Trans. on Automatic Control **50**, 768 (2005).

[28] R. van Handel, J. K. Stockton, and H. Mabuch, IEEE Transactions on Automatic Control **50**, 768 (2005).

[29] S L Adler and D C Brody and T A Brun and L P Hughston, Journal of Physics A: Mathematical and General, **34**, 8795 (2001).

[30] P. Goetsch and R. Graham, Phys. Rev. A. **50**, 5242 (1994).

[31] H. M. Wiseman J. Opt. B: Quantum and Semiclassical Opt. **8** , 205 (1996).

[32] K. Jacobs and P. L. Knight, Phys. Rev. A. **57**, 2301 (1998).

[33] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, (Cambridge University Press, Cambridge, 2000).

[34] R. Bhatia, R., *Matrix Analysis*, (Graduate Texts in Mathematics, 1996).

[35] Wolfram Research, Inc., *Finite summation of Csc*, `http://functions.wolfram.com/01.10.23.0002.01`, accessed 14/04/2009.

[36] J. C. Varilly and J. M. Gracia-Bondia, Annals of physics **190**, 107 (1989).

[37] J. Combes and H. M. Wiseman, J.Opt. B: Quant. and Semiclass. **7**, 14 (2005).

[38] F. T. Arecchi and E. Courtens and R. Gilmore and H. Thomas, Phys. Rev. A **6**, 2211 (1972).

[39] D. T. Pegg and S. M. Barnett, EPL **6**, 483 (1988).

[40] D. T. Pegg and S. M. Barnett, Phys. Rev. A **39**, 1665 (1989).

[41] S. T. Merkel, G. Brennen, P. S. Jessen, and I. H. Deutsch, Phys. Rev. A **80**, 023424 (2009).

[42] W. Schleich, *Quantum optics in phase space*, (Vch Verlagsgesellschaft Mbh, 2001).

[43] A. Klappenecker and M. Rötteler, Proc. Int. Conf. on Finite Fields and Applications, LNCS–springer, **2948**, 137 (2004).

[44] C Hill and J. F. Ralph, New J. Phys., **9**, 151 (2007).

[45] A.B. Klimov, L.L. Sánchez-Soto, and H. Guise, Journal of Physics A: Mathematical and General **38**, 2747 (2005).

[46] J. Lawrence, Č. Brukner and A. Zeilinger, Phys. Rev. A **65**, 032320 (2002).

[47] J. L. Romero, G. Björk, A. B. Klimov and L. L. Sánchez-Soto, Phys. Rev. A **72**, 062310, (2005).

[48] P. Kurzyński, W. Kaszub and M. Czechlewski, Journal of Physics A: Mathematical and Theoretical **43**, 265303 (2010).

[49] S. Massar and P. Spindel, Phys. Rev. Lett. **100**, 190401, (2008).

[50] A. Shabani and K. Jacobs, Phys. Rev. Lett. **101**, 230403 (2008).

[51] K. Jacobs and A. Shabani, Contemporary Physics **49**, 435 (2008).

[52] V. P. Belavkin and A. Negretti and K. Mølmer, Phys. Rev. A, **79**, 022123 (2009).

[53] J. Combes and H. Wiseman, J. Phys. B: At. Mol. Opt. Phys. **44** 154008 (2011).

[54] J. Chiaverini et al., Science **308**, 997 (2005).

[55] J. Petta et al., Science **309**, 2180 (2005).

[56] R. Bianchetti et al., Phys. Rev. Lett. **105**, 223601 (2010).

[57] M. Neeley et al., Science **325**, 722 (2009).

[58] The order and placement of the eigenvalues on the diagonal will depend on the initial state.

[59] P. E. F. Mendonça and J. Combes, unpublished, 2006.

[60] Combes and Jacobs August 2008, unpublished.

[61] In the simulations we used a linear trajectory type simulation so that positivity of the state is ensured.

[62] $M_0 = [[1, 0, 0, 0]', [0, 1, 0, 0]', [0, 0, 1, 0]', [0, 0, 0, 1]']$, $M_1 = [[1, 1, 1, 1]', [1, 1, -1, -1]', [1, -1, -1, 1]', [1, -1, 1, -1]']/2$, $M_2 = [[1, -1, -i, -i]', [1, -1, i, i]', [1, 1, i, -i]', [1, 1, -i, i]']/2$, $M_3 = [[1, -i, 1, i]', [1, i, 1, -i]', [1, -i, -1, -i]', [1, i, -1, i]']/2$, $M_4 = [[1, -i, -i, -1]', [1, -i, i, 1]', [1, i, i, -1]', [1, i, -i, 1]']/2$.